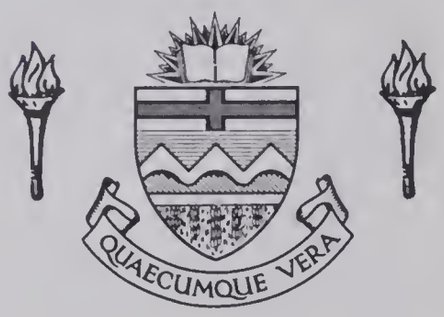

T H E U N I V E R S I T Y O F A L B E R T A

RELEASE FORM

NAME OF AUTHOR Jerry Tadeusz Broszkowski

TITLE OF THESIS Meteorological Applications of Empirical Orthogonal Functions

DEGREE FOR WHICH THESIS WAS PRESENTED Master of Science

YEAR THIS DEGREE GRANTED 1979

THE UNIVERSITY OF ALBERTA

METEOROLOGICAL APPLICATIONS OF EMPIRICAL ORTHOGONAL FUNCTIONS

by

© JERRY TADEUSZ BROSZKOWSKI

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

IN

METEOROLOGY

DEPARTMENT OF GEOGRAPHY

EDMONTON, ALBERTA

FALL, 1979

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled "Meteorological Applications of Empirical Orthogonal Functions" submitted by Jerry Tadeusz Broszkowski in partial fulfilment of the requirements for the degree of Master of Science in Meteorology.

Dedicated to

Theresa

ABSTRACT

The theory of empirical orthogonal functions is presented by means of a simple two-variable example which can be easily visualized and which can be easily generalized to three or more variables. Applications of the functions are shown by reviewing the available literature with the last two sections introducing complex eigenvector analysis.

Eigenanalysis of a simple wave pattern is followed by the analysis of ten years of monthly mean temperatures in western Canada. The first four time functions of the temperature data are studied by means of periodograms and the use of empirical orthogonal functions for the storage of climatological data is discussed.

500-mb heights in western Canada and the northwestern United States are analyzed next with a practical application being presented: the detection of errors in a set of data. Precipitation data for the summer months in Alberta are also attempted but with discouraging results.

Comments and consideration for future work, such as micrometeorology and complex eigenvector analysis, conclude the work.

ACKNOWLEDGEMENTS

I am grateful to Dr. E.R. Reinelt for his patience and supervision.

I wish to thank Dr. E.P. Lozowski and Dr. K. Smillie for being on my committee.

Thanks also to Mr. P. Kociuba for some of the 500-mb data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Empirical orthogonal functions or principal components arise in the statistical analysis of one or more variables by a method that can be described as a translation plus a rotation of axes or, alternatively, as a representation of the data by a set of orthogonal* functions.

An example of the method of representation by orthogonal functions is shown in Appendix I; it is sufficient to say now that if there are N stations observing some variable p at several times t, then p at station j and time t may be represented by the expansion:

$$p_j(t) = \sum_{i=1}^{N} q_i(t) y_{ij}$$

where the $q_i(t)$ are time functions and the y_{ij} are the empirical orthogonal functions. Of course, the y_{ij} could be any of the standard orthogonal polynomials (Legendre, Tschebysheff, etc.) but it is interesting to see what happens when the functions are derived from the observed data. The empirical orthogonal functions can of course be transformed to any other set of orthogonal functions.

---

* To be precise, the functions are orthonormal but will be called orthogonal to be consistent with the literature.

1.2 An Example

Since the application of eigenvector analysis to a complicated set of real data may not provide much insight into what is actually being done, a simple example using translation and rotation of axes will perhaps give a 'feel' for empirical orthogonal functions.

Consider two stations recording temperatures with the data shown in Table 1 and Figure 1 for four consecutive days. Is it possible to reduce the number of variables from two to one and yet retain all the vital information? Also, is it possible to make a reasonable guess as to what the temperature will be on day 4?

In order to reduce the number of variables, one temperature must be a function of the other. One of the simplest relationships is a linear one, so if the two variables correlate linearly to a high degree, one can be written as a linear function of the other. A simple cross-correlation coefficient can then be calculated and this will provide some clue. However, much more information can be obtained if the axes are translated to the 'centre of gravity' of the data and then rotated in such a way that the data correlate highly* along one axis and very little along the other. The two new variables thus produced should be uncorrelated, since if one correlates with the other then one can be written as a linear function of the other, and another transformation may be performed.

The data, with the means removed, are listed in Table 2 and plotted in Figure 2.

---

* In practice, the correlation coefficient is maximized.

Day	Temperature at Station One	Temperature at Station Two
0	9	11
1	21	19
2	29	31
3	21	19

Table 1. Temperatures at two hypothetical stations.

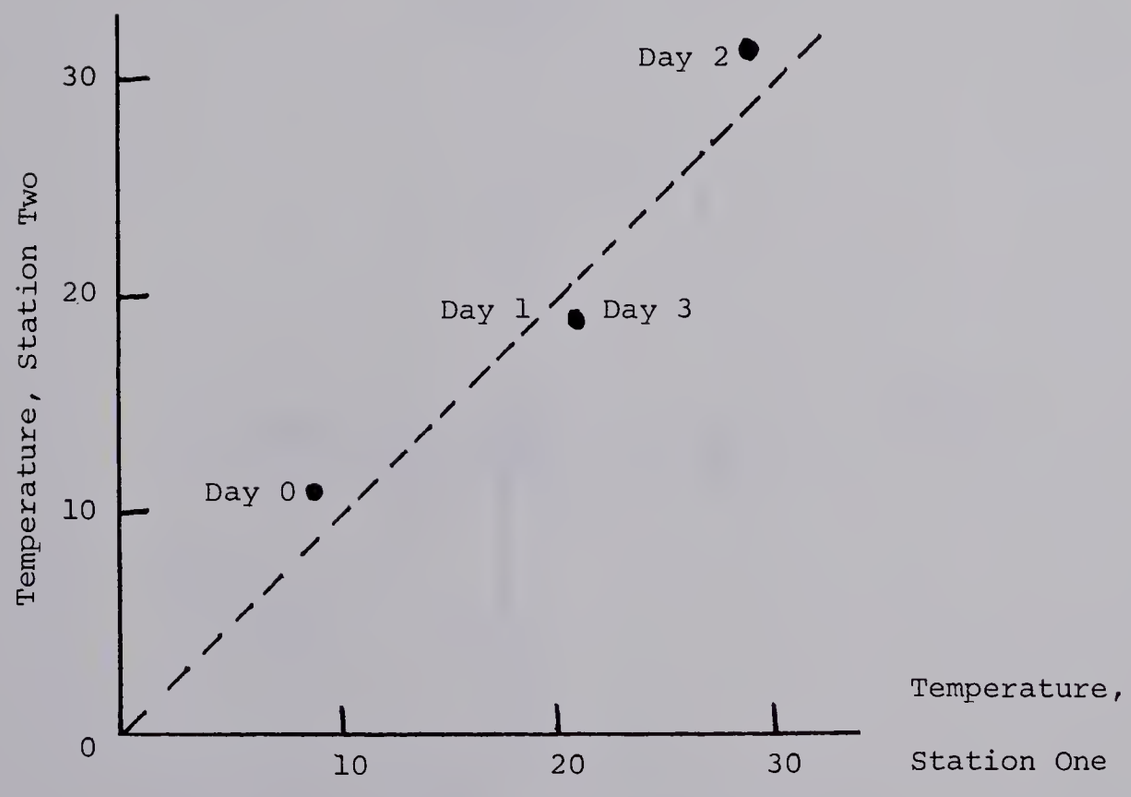

Figure 1. Graph of temperatures from Table 1.

Day	Station One	Station Two
0	-11	-9
1	1	-1
2	9	11
3	1	-1

Table 2. Temperatures from Table 1 with means removed.

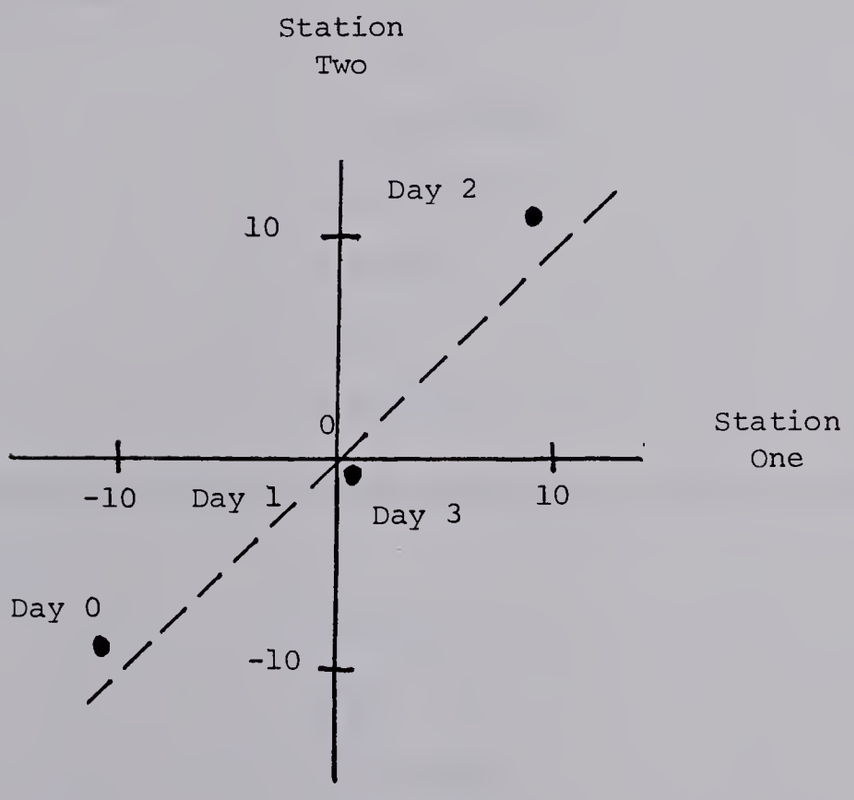

Figure 2. Graph of values from Table 2.

Let the temperatures of Table 1 be represented by the matrix P. By removing the means, let this become P* (i.e. Table 2). By application of a rotation matrix X these values become matrix Q*.

Then, $$Q^* = P^*X \qquad ...1$$

Since X is a rotation matrix,

$$XX^t = I \qquad ...2$$

where I is the identity matrix and X^t is the transpose of X.

Also, the new variables are uncorrelated, so

$$Q^{*t}Q^* = D \qquad ...3$$

where D is a diagonal matrix.

If $$A = P^{*t}P^*$$

then from 1 and 3,

$$D = Q^{*t}Q^*$$

$$= (P^*X)^tP^*X$$

$$= X^tP^{*t}P^*X$$

$$D = X^tAX \qquad ...4$$

Solving $$XX^t = I$$

and $$D = X^tAX$$

for a given A is the well-known eigenvalue problem and can be written as:

$$(A-D)X = 0 \qquad ...5$$

since $$X^tAX = D$$

then $$XX^tAX = XD$$

and $$(AX-DX) = 0$$

For equation 5 to be true,

$$\det |A-D| = 0$$

since $$\det |X| = 1$$

Now, in the present instance

$$A = \begin{pmatrix} 204 & 196 \\ 196 & 204 \end{pmatrix}$$

and the eigenvalues of A are 8 and 400. The corresponding two eigenvectors are a(1,-1) and b(1,1) where a and b are any two scalars. These two vectors form the matrix X:

$$X = a\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

From equation 2, $XX^t = I$

so $$a^2 \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and $$a = \frac{\sqrt{2}}{2} = .71$$

so X must be:

$$X = \begin{pmatrix} .71 & .71 \\ .71 & -.71 \end{pmatrix}$$

If the axes are rotated counterclockwise through 45° so that one passes through (.71,.71) while the other passes through (.71,-.71) then the new variables become by equation 1:

$$Q^* = \begin{pmatrix} -11 & -9 \\ 1 & -1 \\ 9 & 11 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} .71 & .71 \\ .71 & -.71 \end{pmatrix} = \begin{pmatrix} -14.2 & -1.42 \\ 0.0 & 1.42 \\ 14.2 & -1.42 \\ 0.0 & 1.42 \end{pmatrix}$$

From equation 1 it can be seen that,

$$P^* = Q^*X^t$$

or

$$\begin{pmatrix} -11 & -9 \\ 1 & -1 \\ 9 & 11 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} -14.2 & -1.42 \\ 0 & 1.42 \\ 14.2 & -1.42 \\ 0 & 1.42 \end{pmatrix} \begin{pmatrix} .71 & .71 \\ .71 & -.71 \end{pmatrix}$$

$$= \begin{pmatrix} -10.0-\underline{1.0} & -10.0+\underline{1.0} \\ 0.0+\underline{1.0} & 0.0-\underline{1.0} \\ 10.0-\underline{1.0} & 10.0+\underline{1.0} \\ 0.0+\underline{1.0} & 0.0-\underline{1.0} \end{pmatrix}$$

If the original temperatures were read to an accuracy of 1° or if an accuracy of 1° is acceptable, then only the first eigenvector with its coefficients plus the means can be used to represent all the data of Table 1.

i.e.

$$P^* \doteq \begin{pmatrix} -14.2 \\ 0 \\ 14.2 \\ 0 \end{pmatrix} (.71 \quad .71) = \begin{pmatrix} -10 & -10 \\ 0 & 0 \\ 10 & 10 \\ 0 & 0 \end{pmatrix}$$

and

$$P \doteq \begin{pmatrix} -10 & -10 \\ 0 & 0 \\ 10 & 10 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 20 & 20 \\ 20 & 20 \\ 20 & 20 \\ 20 & 20 \end{pmatrix} = \begin{pmatrix} 10 & 10 \\ 20 & 20 \\ 30 & 30 \\ 20 & 20 \end{pmatrix}$$

Hence the number of variables can be reduced if the slight errors that result are assumed to be noise and not significant values. The representation is in fact very good since it can be shown that over 98% of the variance of the data is explained by

the first eigenvector.

Can the next day's temperature be predicted? If the eigenvectors are thought of as the spatial components of the temperature then the coefficients can be thought of as the time component. If N_d is the day number, then based on four days' data, the first time function can be written as $-14.2\cos(N_d\pi/2)$, which has a period of 4 days, and the second as $-1.42\cos(N_d\pi)$ which has a period of 2 days. Day four's temperature would then be expected to be (10,10) using only the first eigenvector and its associated time function.

There is no need to restrict the number of stations to two or the number of different variables to one. The number of stations and variables is limited only by the ability to find the eigenvalues and eigenvectors of a square matrix with dimension equal to the number of stations times the number of variables. The data need not be fitted to any sort of grid, and the spacing of the data points depends on the scale of the features being studied. The time intervals also need not be equal. With each station can be associated any number of parameters and if the units of the parameters differ it is best to normalize them before analysis.

If there are 50 stations measuring 3 variables each then the eigenvalues of a 150-square matrix will have to be found. Though no easy feat, with today's and tomorrow's computers and modern numerical methods even a matrix of this size can be handled fairly efficiently.

CHAPTER 2

REVIEW OF THE LITERATURE

2.1 Lorenz, E.N. (1956)

Lorenz's paper of 1956 was one of the first to show how and why empirical orthogonal functions can be used in meteorology. He looked at the problems of representing a predictand, $x(t)$, by a set of predictors $p_1(t),\ldots,p_M(t)$, where t may be time or any other variable, by using the method of least squares, and pointed out that prediction formulae derived from different samples will not in general be similar, and that any one of them will not necessarily be the best formula for the whole population. There are the dangers that a formula good for one sample may be poor for the population, and a formula that is not the best one for a sample may be the best one for the population, dangers which are inescapable unless the whole population is available. If the entire population is not available then the number of predictors should be limited. By using statistical and practical results Lorenz demonstrated the need to keep the ratio of the number of predictors to the number of observations low.

As a method of reducing the number of predictors, Lorenz derived the representation of them by empirical orthogonal functions, and pointed out that this method is analogous to factor analysis as used by other disciplines, psychology in particular.

As a test, sea-level pressures from 64 stations in southern Canada and the United States, as observed at 1230Z each day in February from 1949 to 1953 were used to compute the first 16 eigenvectors. It was found that the first 8 functions specified 91% of the variance and that 16 specified 97%. When the first 8 were used to calculate the time coefficients for an independent set of data (the years 1947 and 1948), it was found that again 91% of the mean square error could be explained.

When the eigenvectors were drawn up, one particular note was made that is observed in almost all analyses by means of empirical orthogonal functions: the lower-numbered eigenvectors have large wavelengths and represent large-scale features while higher-numbered eigenvectors have short wavelengths and represent smaller-scale features or perhaps only noise.

Lorenz next attempted to predict the sea-level pressure field of the 64 stations by using the field 24 hours earlier as a predictor. This was accomplished by using the first K eigenvectors to represent the pressure field on day i, calculating the first J time functions for day i+1 (J=1 to 8) for values of K=1 to 8 and then combining the first J time functions with the first K eigenvectors to get the pressure field on day i+1. Predictions for both the development sample and the independent sample were attempted. For the development sample a maximum reduction of variance of 50% resulted from using 8 predictors and 8 predictands, while for the independent sample a reduction of 31% was seen for J=4 and K=5. For the 8 and 8 combination the reduction was 30%.

Although the results were not especially encouraging, Lorenz pointed out that this was only a feasibility study and better results most certainly would have been obtained if the two samples had been larger, if predictors for stations on the boundary had not been included in the calculation of reduction of error, and if a better scheme using more variables had been used.

In conclusion Lorenz suggested that empirical orthogonal functions could have more uses than just representing a series of predictors by a smaller set. Classifying meteorological phenomena, simplifying non-linear statistical prediction and perhaps offering a tool in dynamical prediction were felt by the author to have promising applications.

2.2 Gilman, D.L. (1957)

Utilizing monthly mean sea-level pressures and monthly mean surface temperatures, Gilman's objectives were: a reduction in the number of parameters needed to specify the monthly mean circulation in the northern hemisphere and the monthly mean surface temperature in the United States, an auto-correlation and cross-correlation of the two variables at lags zero and one month, and a study of the linear regression formulae derived from the correlations.

After proving some results from least-square fitting and orthogonal functions, and discussing some other researchers' work with orthogonal functions, the author derives the method of empirical orthogonal functions in two ways, mentions some nomenclature and points out Lorenz's proof that empirical orthogonal functions provide an optimal representation of data as compared to any other

type of function.

From the Northern Hemisphere Historical Map Series, 1899-1939, monthly mean sea-level pressures and temperatures were obtained for the months December, January and February, and a test sample was acquired for the months December to March, 1947 to 1956. The pressure data were treated first.

The initial grid consisted of points separated 10° latitudinally, and 20° longitudinally from 20°N to 60°N; 40° longitudinally from 70°N to 80°N. Analysis of some daily maps in the Historical Series, however, showed that the data in some areas were suspect and therefore the number of grid points was reduced to 90. The computer available to Gilman in 1957 was insufficient to diagonalize a 90^{th} order matrix, so a method was used where the first 32 eigenvectors could be generated by first partitioning the correlation matrix. The means were first removed from the pressure field "in order to spare the system of orthogonal functions the trouble of having always to reconstruct the normal before being able to proceed to the abnormal." This was done by subtracting the 10-year December means from the mean monthly December pressures, the January means from the January pressures, and the February means from the February pressures. Then the pressure deviations were divided by the standard deviations to produce standardized variables, under the assumption that the variances of monthly mean pressures remained nearly unchanged throughout the winter. This was performed because the highest value of the variance was approximately thirty-times greater than the lowest value.

The 32 eigenvectors obtained from the standardized pressures accounted for 90.1% of the variance with the first contributing only 12.5%. (See Figure 3 for a reproduction of pressure eigenvector number 1.) Gilman comments that, on the basis of plots of the first 8 eigenvectors, the scale of the pattern decreases as the order of the eigenvectors increases, with the patterns becoming more cellular, and also that "... they tend to emphasize the classical 'centers of action', Iceland, Alaska and the Aleutians, the Central Pacific, Central Asia, the Azores - in spite of the even initial weighting of all grid points..."

Gilman points out that prediction equations for temperature, or other atmospheric variables, will be more stable if the variables are represented by empirical orthogonal functions, since the unpredictable components, such as small-scale effects and noise, can be filtered out. Using 30 stations in the United States (shown in Figure 4) he then proceeded to standardize the monthly mean temperatures and generate the 30 eigenvectors and their time coefficients.

The first function contributed 38.3% of the variance while the first three increased this to 81.1%, and six accounted for over 90%. Gilman concluded that for the regression equations to be developed in this study, the first three eigenvectors would suffice, since the higher-numbered eigenvectors contribute only small amounts to the total variance, and may or may not be predictable.

The author now endeavored to formulate regression equations using up to 20 leading time coefficients of the standardized monthly mean pressure field as predictors (20 eigenvectors of the pressure field specified 84.6% of the variance) and the first six time

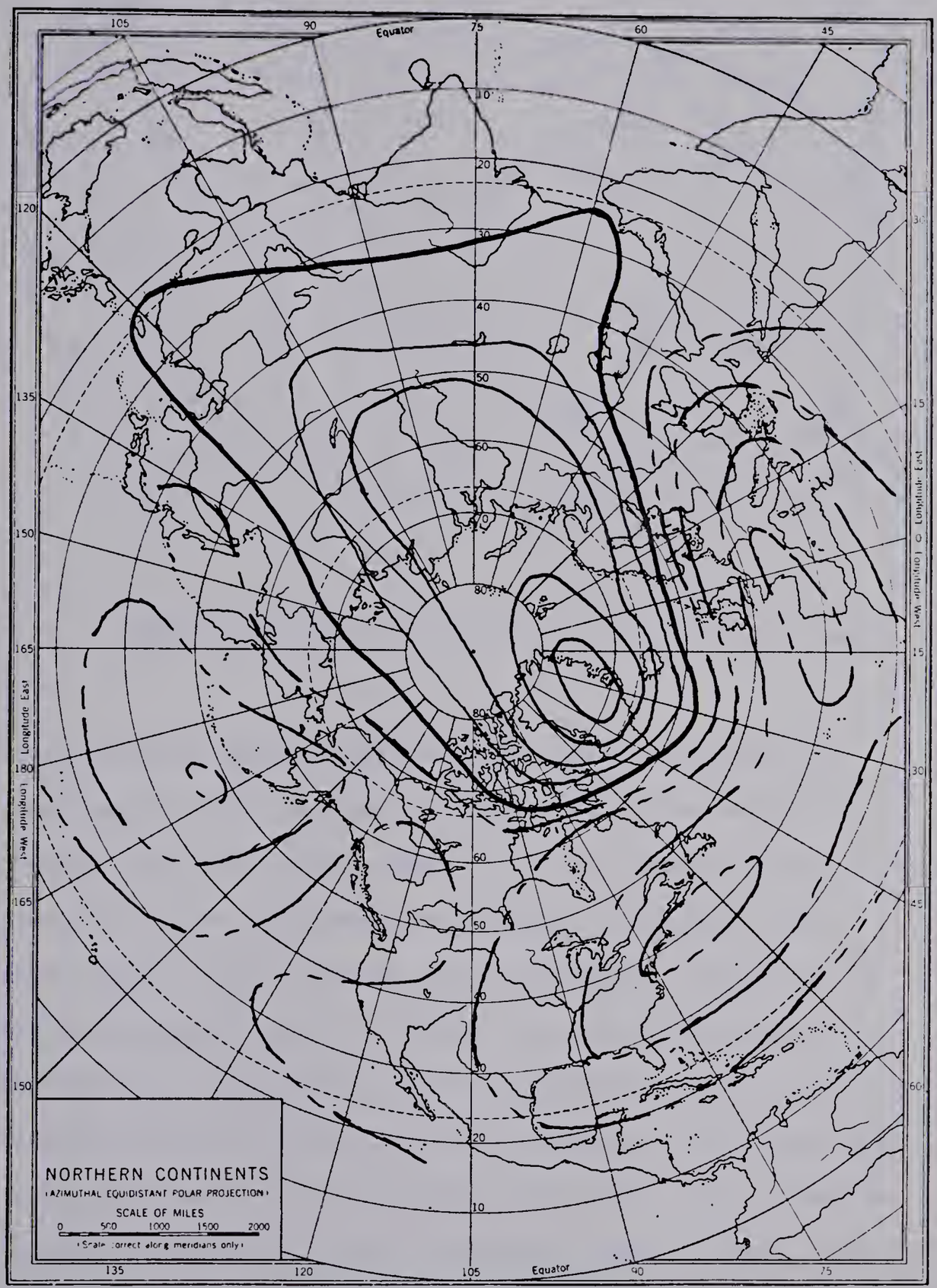

Figure 3. Hemispheric function I. (After Gilman, 1957)

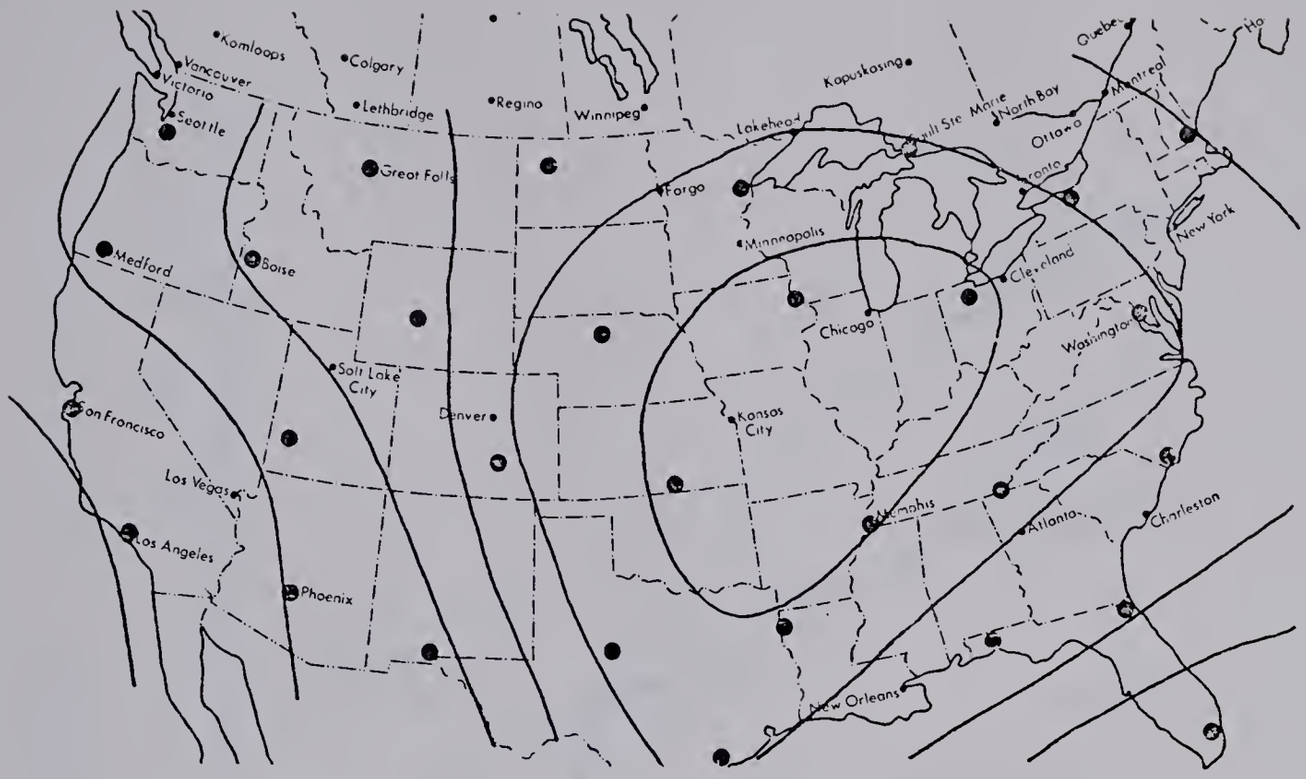

Figure 4. Temperature function I. (After Gilman, 1957)

coefficients of the standardized monthly mean temperatures as predictands for the same month as the pressure time coefficients, and then one month later. Prediction of the pressure time coefficients at lag one month was also attempted. Figures 5, 6 and 7 show Gilman's results for the development sample, as well as the values expected by chance. He points out that the variation in results was not surprising, but the explanation of the good specification results in physical terms was difficult. The author suggested that one possible explanation was "simple linear advection to and from the climatological sources and sinks of heat"; and another plausible cause was large-scale horizontal eddy transport in the monthly mean circulation pattern. The reasons for the poor

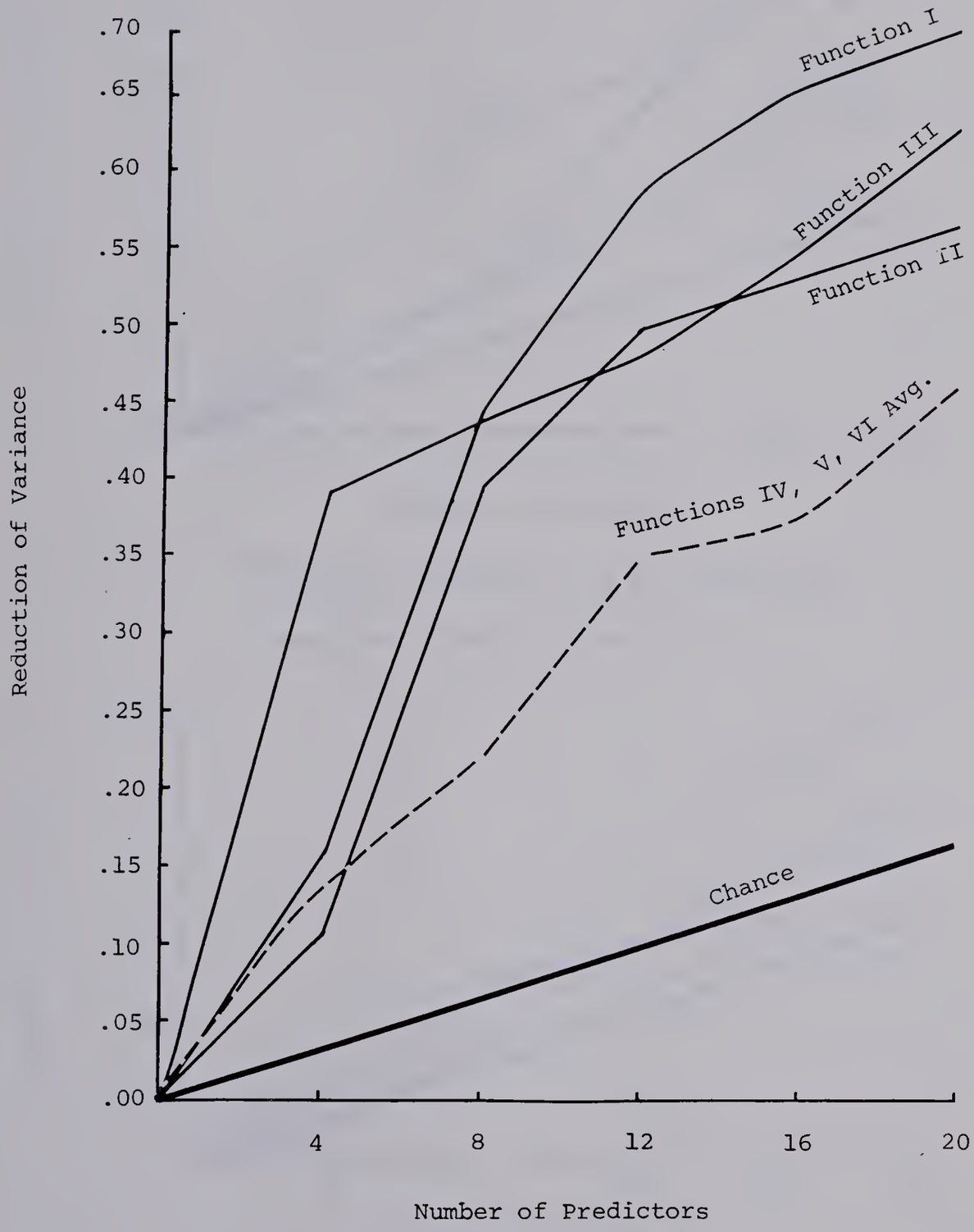

Figure 5. Specification of the temperature functions. (After Gilman, 1957)

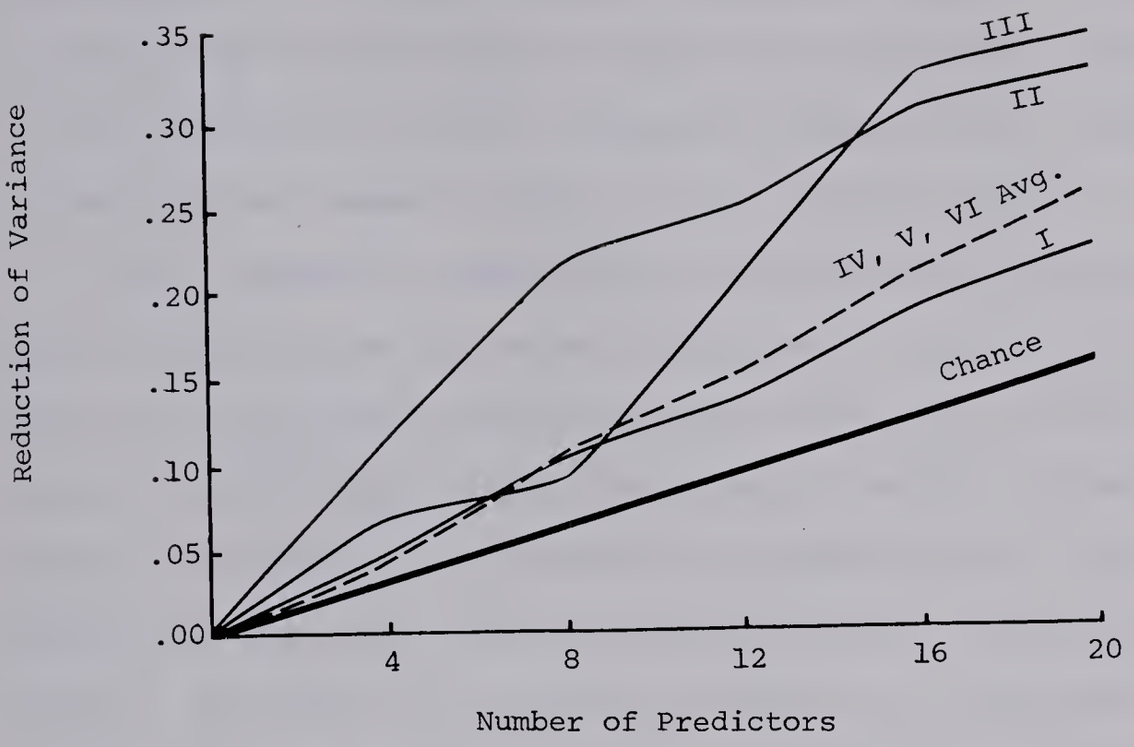

Figure 6. Prediction of the temperature functions. (After Gilman, 1957)

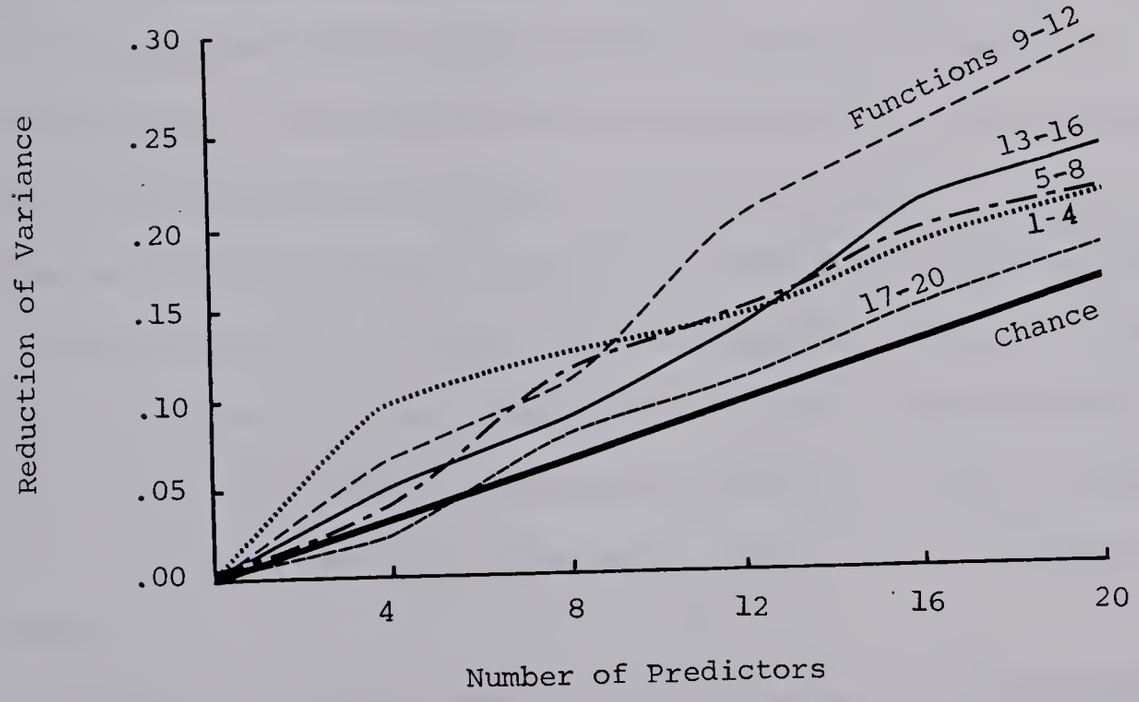

Figure 7. Prediction of the pressure functions. (After Gilman, 1957)

predictability of the time functions as well as the small variation of predictability with respect to time-function number was explained by the "complexity of the full hemispheric monthly pressure pattern and the relative freedom of action of their different limbs..."

After digressing on prediction by climatology, persistence and chance, Gilman applied the regression equations to the test data of 1947 to 1956, and made comparisons between these results and the results from the aforementioned prognostication methods. As before, the time functions for the pressures and temperatures were first obtained, although this time unstandardized pressure anomalies were employed. The results are presented in Figure 8 for temperature function No. 2, the one best predicted. Time functions No. 4 and higher showed unpromising results. Interestingly, the second temperature eigenvector in the test sample accounted for about 40% of the variance, while the first and third explained 27% and 10%, respectively. The equations for the specification of the temperature functions were also applied, with the outcome that although there was little difference between the results from the development and test samples for the first three temperature functions, numbers four to six were well below the expected values. Gilman comments that "the optimum number of pressure functions with which to specify the national temperature field again seems to lie...somewhere between 12 and 16."

Gilman then reconstructed the station values of the temperature from the first three eigenvectors. Calculating and plotting the

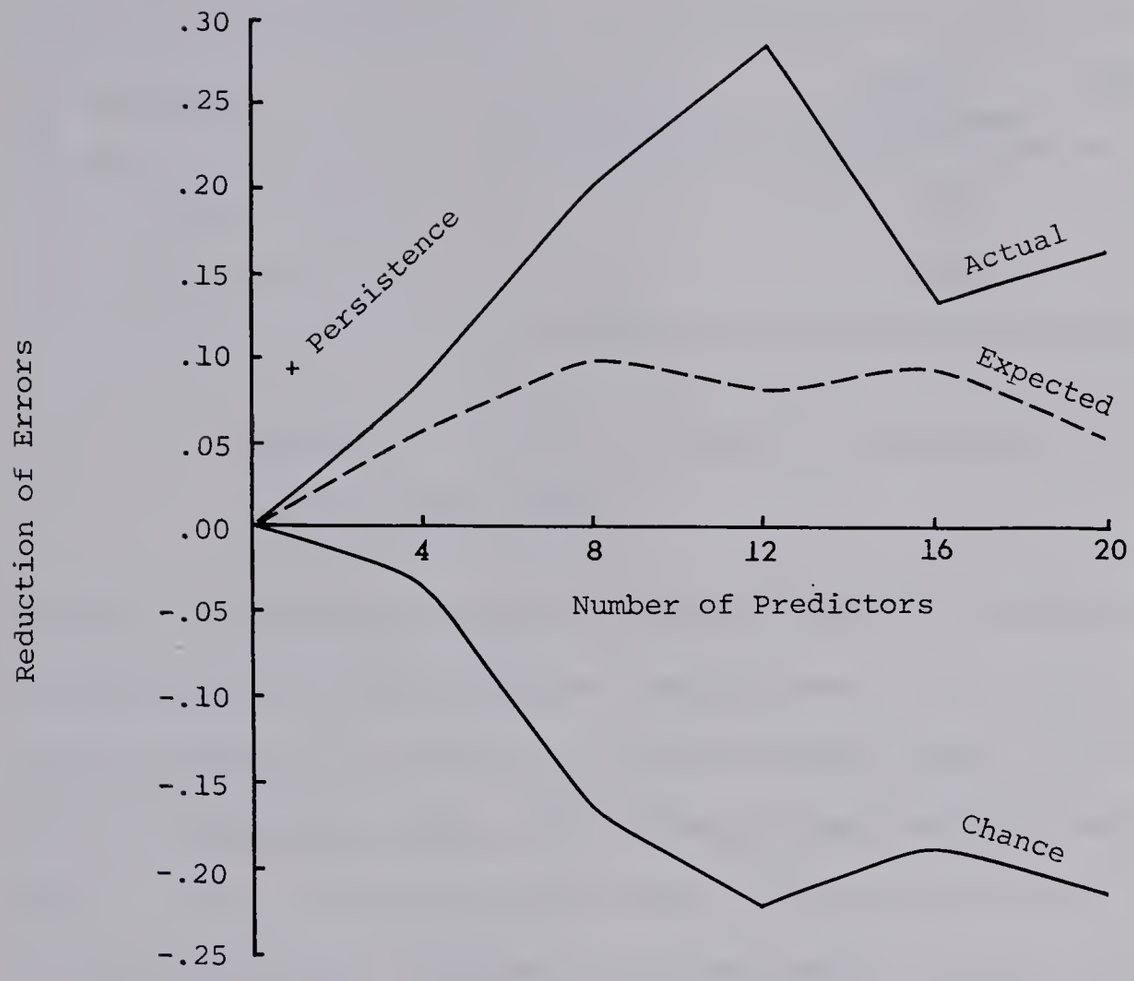

Figure 8. Prediction of temperature function II. (After Gilman, 1957)

reduction in variance or error for the thirty stations, he concluded, because of the similarity between the independent and dependent samples, that the temperature eigenvectors "maintain almost perfect mutual orthogonality when applied to the independent data." Gilman also surmised that the synoptic conditions represented by the different eigenvectors have "largely independent physical causes."

Utilizing three temperature time functions specified and predicted by the regression equations, and using 16 pressure functions for specification and 12 for prediction, reductions of error or

Predictor	12 Empirical functions	Reduced Persistence	Chance
Dependent sample	+.14	+.04	- - -
Independent sample	+.12	+.035	-.13

Table 3. Prediction of station temperatures.

(After Gilman, 1957)

variance were again calculated and plotted for the 30 stations for the development and test samples. For the specification map, the average reduction of error was 0.43 for the test sample, and 0.48 for the development sample. The values for the prediction data are shown in Table 3; all values are "above" climatology. Gilman remarks on the results that: "because of the very modest, though positive, levels of verification obtained in this work, a serious attempt at physical interpretation is probably not justified."

Since thirty-day forecasts of the United States Weather Bureau seemed to have about the same consistency as persistence, which is almost always negative with respect to climatology, Gilman concluded that the regression equations from empirical functions are probably better thirty-day predictors.

2.3 Craddock, J.M. and Flood, C.R. (1969)

Craddock and Flood attempted "to reduce the raw material of the long-range forecaster to manageable proportions by removing redundancies, and representing important fields in terms of the

smallest possible number of mutually independent variables." Using the 500-mb field over the Northern Hemisphere, they hoped not only to reduce the amount of data by using orthogonal functions but also to filter out most of the noise.

The actual data consisted of daily 500-mb heights at 130 stations scattered throughout the Northern Hemisphere north of 30°N for the years 1965 to 1967. It could have been possible to use 540 points and the years 1949 to 1967, but this was not feasible because of too much data lacking at some points and the shortcomings of the available computer. A 106-point grid however, was processed. Even with 130 points many values were missing, especially over the Pacific, and these were replaced by values observed one day earlier.

Craddock and Flood computed the eigenvectors of the 50 largest eigenvalues and found that 97.1% of the variance was explained by these 50 eigenvectors. The matrix of 50 x 1095 time coefficients was also calculated, with the mean, standard deviation, coefficients of skewness and kurtosis, and the first 60 serial correlations being calculated for each series of 365 coefficients corresponding to each year and each eigenvalue.

Craddock and Flood now examined the patterns formed by the main eigenvectors with notes on some of the time functions. The easiest to explain was the first pattern (see Figure 9) which was very similar to the map of the total variance of the 500-mb field, had no strong gradients and a time function which closely resembled a sine wave of period one year. The rest of the eigenvector diagrams

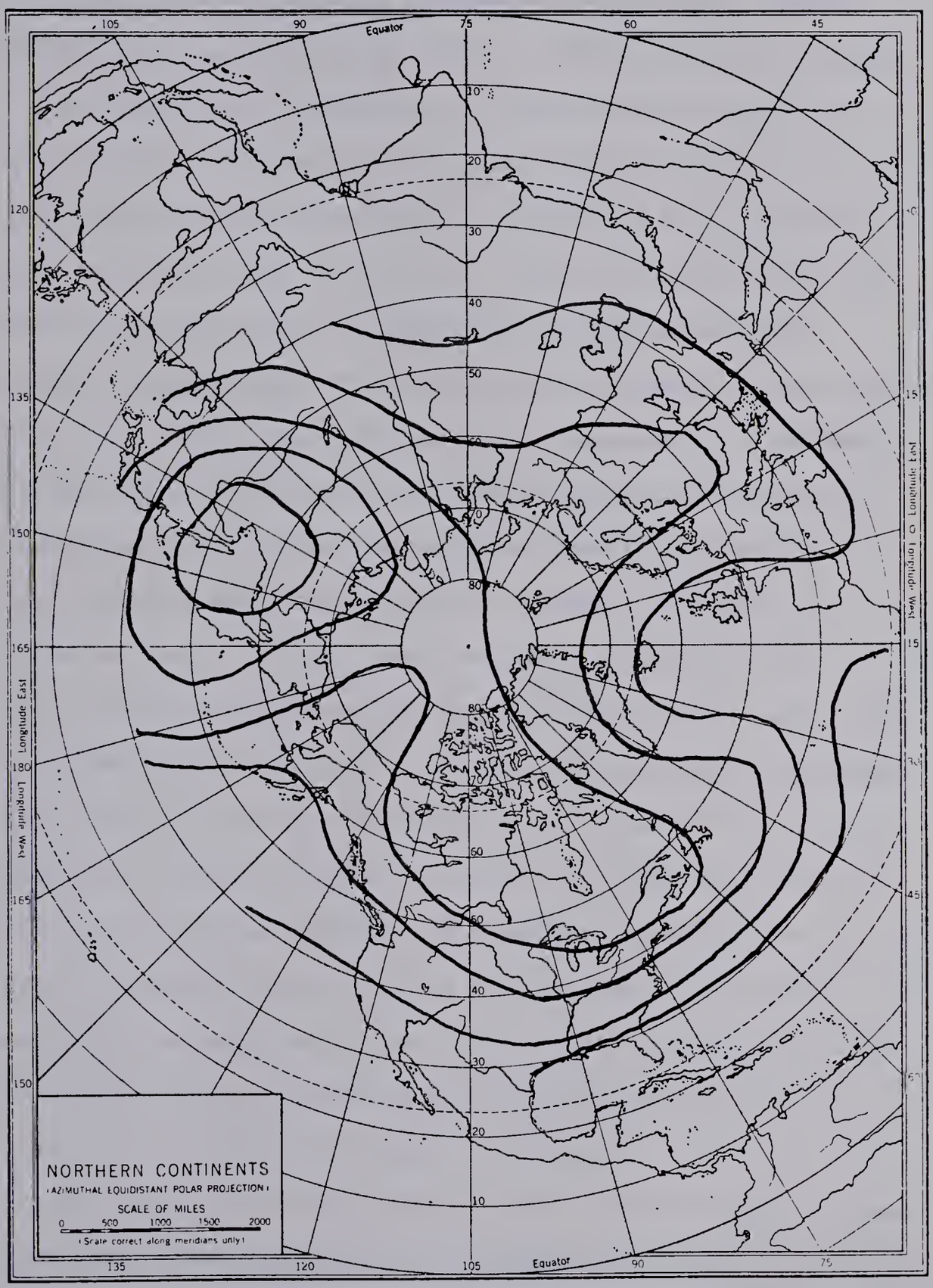

Figure 9. Terms of Eigenvector No. 1. (After Craddock and Flood, 1969)

showed progressively stronger gradients with no readily explainable patterns, and time series with no visibly interpretable graphs.

The authors next considered the problem of the optimum representation of the data by the minimum number of eigenvectors without any loss of "meteorological content" and still filtering out a substantial part of the noise. Since 15 eigenvectors represented 85% of the total variance while 50 represented 97.1%, Craddock and Flood suggested that 15 would be enough for some purposes such as forecasting, while 50 would certainly be a maximum. Although their next argument, that the difference between the recombination of 50 eigenvectors and their coefficients and the original data will be within the errors of chart analysis and the reading of values, is true, their procedure is perhaps faulty. "When a similar analysis was carried out on data for the same grid, but excluding the Pacific area, the residual variance not accounted for by the first 50 eigenvectors was only 1.1 per cent, so that the 2.9 per cent observed here must largely arise from errors in estimating missing values in the Pacific sector." Definitely some of the errors derived from estimating values contributed towards the unexplained variance, but if the Pacific area had 80 stations, then 100% of the variance would have been explained by the first 50 eigenvectors, regardless of errors in estimation.

The logarithms of the eigenvalues were plotted versus the eigenvector numbers as further proof that eigenvectors above approximately number 50 were principally noise. (Figure 10) The authors noted that after roughly 50 eigenvectors, the logarithms of the eigenvalues

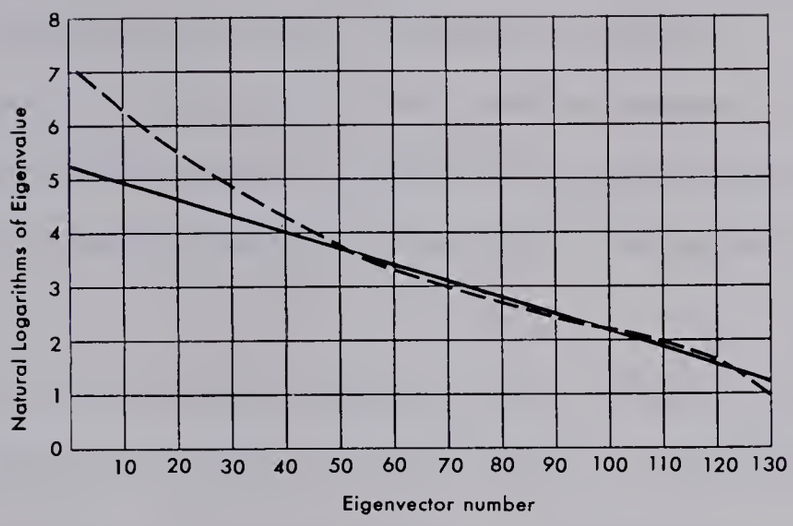

Figure 10. The relation between eigenvector number and the natural logarithm of the corresponding eigenvalue. (After Craddock and Flood, 1969)

closely fit a straight line and hence, "...although the theoretical justification is rather obscure, our work afford practical support for the statement 'In meteorology noise eigenvalues are in geometric progression'". Thus the conclusion that the first 20 to 25 eigenvectors are almost all true data and that the ones greater than 40 are almost all noise.

After transforming the autocorrelation coefficients at a lag of one day for each series of time coefficients into Fisher's Z-statistic* and averaging each Z over the three years, Craddock and Flood found that the Z-values dropped fairly rapidly and were

---

* $Z = \frac{1}{2}\log_e \frac{(1 + r)}{(1 - r)}$, where r is the correlation coefficient and is distributed normally with a variance of $1/(N-3)$, where N is the number of observations.

extrapolated to be almost zero after eigenvector number 60. Since a low Z value implies a low correlation and hence no harmonics of the annual variation, they concluded that about 60% of the total variance was due to basically annual variations and 40% to non-periodic processes.

The quality of the data was examined by two methods: calculation of the coefficients of kurtosis of the eigenvector coefficients, and by comparing the original field of January 2, 1965 with the recombined field using eigenvectors from 1 to 49. The first method was applied to the coefficients from the 106-point grid and the years 1949 to 1967. An impossible value in the coefficient of kurtosis resulted for a number of charts and the authors found major errors in those charts. The number of extreme coefficients decreased with the years, however, and none were found in the 1965-1967 data.

Using the comparison method, when 35 eigenvectors were used, two points in the Pacific Ocean were shown to be wrong although the errors were within three standard deviations and might have been hard to find by other means. If only tables of 500-mb heights were available and not charts, it may have been difficult to reject these values. Craddock and Flood remark: "The question of deciding whether any such feature is rare but genuine, or as usual due to an incorrect observation, is a matter for human judgment in each individual case."

After some comparisons with other work the authors state their main conclusion: "the number of degrees of freedom of the planetary

airflow over the Northern Hemisphere is somewhat less than 50."

2.4 Craddock, J.M. and Flintoff, S. (1970)

Using the same grid and the same time period as Craddock and Flood (1969), Craddock and Flintoff analyzed with empirical orthogonal functions the 1000-mb heights and the 1000-500-mb thickness lines with the objective of comparing the resultant eigenvectors with those from the 500-mb field as determined by Craddock and Flood.

Although the first eigenvector of the 1000-mb field accounted for only 21.6% of the variance, the variance attributed to the first 50 eigenvectors totaled 93.7%. The figures for the thickness lines were 73.1% and 96.8%, respectively. A plot of the logarithms of the eigenvalues of the two sets of data was presented with the values of Figure 10 also shown. Since the two new graphs likewise became linear after approximately eigenvector number 50, the authors concluded that as in the 500-mb case, "... the eigenvectors which represent nothing but noise are those numbered from about 50 upwards."

Noting that the contours of the first eigenvector of the 500-mb heights and the first eigenvector of the thickness lines were similar, and that those of the second 500-mb, second 1000-mb and seventh thickness patterns were similar, Craddock and Flintoff investigated the possibility of representing one set of eigenvectors in terms of another. Since any eigenvector in one set can be completely specified by 130 eigenvectors of another set, the authors tried to determine how well the first 50 thickness eigenvectors and

the first 50 1000-mb eigenvectors could be generated from the first 20 or the first 50 eigenvectors of the 500-mb analysis.

Although the first 20 eigenvectors were not very satisfactory in representing any thickness or 1000-mb eigenvectors numbered higher than about 15, the first 50 eigenvectors of the 500-mb field gave convincing results for the first 45 eigenvectors in both the other fields. Table 4 summarizes the remainder of the authors' results.

Craddock and Flintoff conclude that the use of the 50 500-mb eigenvectors is very efficient in the sense that 50 eigenvectors can more accurately generate three different fields than can the three best sets of 35 eigenvectors. Emphasized also was the advantage of using one basis for three separate fields when all three were being studied simultaneously.

Data	500-mb Heights	1000-500-mb Thickness Lines	1000-500-mb Thickness Lines	1000-mb Heights	1000-mb Heights
Represented by 50 Eigenvectors from:	500-mb Heights	Thickness Lines	500-mb Heights	1000-mb Heights	500-mb Heights
% Variance Explained	97.1	96.8	95.8	93.7	89.4

Table 4. Comparison of variables represented by eigenvectors generated from the variables themselves and from other variables.

2.5 Kutzbach, J.E. (1967)

Kutzbach's objective was the study of the combined eigenvector representation of monthly mean sea-level pressure, surface temperature and precipitation in 23 regions of North America for 25 Januaries and "an examination of their (the combined representation's) synoptic consistency." After a review of previous work he used the method of Lagrange multipliers to derive the theory of empirical orthogonal functions. Kutzbach also gave a geometrical interpretation of the eigenvectors derived from a set of data.

Since only 70 by 70 matrices could be diagonalized by the then available techniques, a maximum of 23 points could be used if the required three variables were to be associated with each point. The points were chosen to be near the centers of 23 roughly equal-area regions out of a total of 46 covering North America. (The points are distinguishable in Figure 11.) The monthly mean sea-level pressure, surface temperature and precipitation values assigned to each point were from the Januaries of 1941 to 1965 and were the average of 2 to 5 climatological stations within each region. The variables were normalized, not only to weight each variable equally, but also because the normalized fields of temperature and precipitation resemble the climatological classifications of below normal, normal, etc. more closely than the departure fields of those two variables.

Kutzbach generated eigenvectors of pressure, which he denoted as (P), temperature (T), precipitation (R), pressure and temperature (PT)

and pressure, temperature and precipitation (PTR). After pointing out similarities between his T-eigenvectors and the temperature eigenvectors of Gilman (1957) and remarking on the similarities of the combined and separate representations, Kutzbach commented on the "synoptic consistency of the departure patterns of the climatic variables in the combined eigenvector representations."

Figure 11 shows the first eigenvectors of PT while Figure 12, those of PTR, where the eigenvectors of P are solid lines, of T dashed lines and of R dashed-dotted lines. One "synoptic consistency" is easily discernible in Figure 11. Kutzbach noted that with a more northerly than "normal" flow from the north-west, a region of negative temperature departures occurs in southwestern Canada and the northwestern United States. Other interrelationships are discussed as well. The author comments that, in general, the interrelationships between the first five eigenvectors of PTR could be explained synoptically.

Kutzbach also examined the variances explained by the different eigenvectors, the limitations of the patterns and the time coefficients of the eigenvectors.

The cumulative variances of the different eigenvector representations are shown in Table 5. The author noted that in order to explain a specified amount of variance in a combined representation, PT or PTR, fewer eigenvectors were needed than the total number of separate eigenvectors required to explain the same amount of variance, and thus he concluded that the combined representation is more efficient. Kutzbach also commented that "for a given set of M

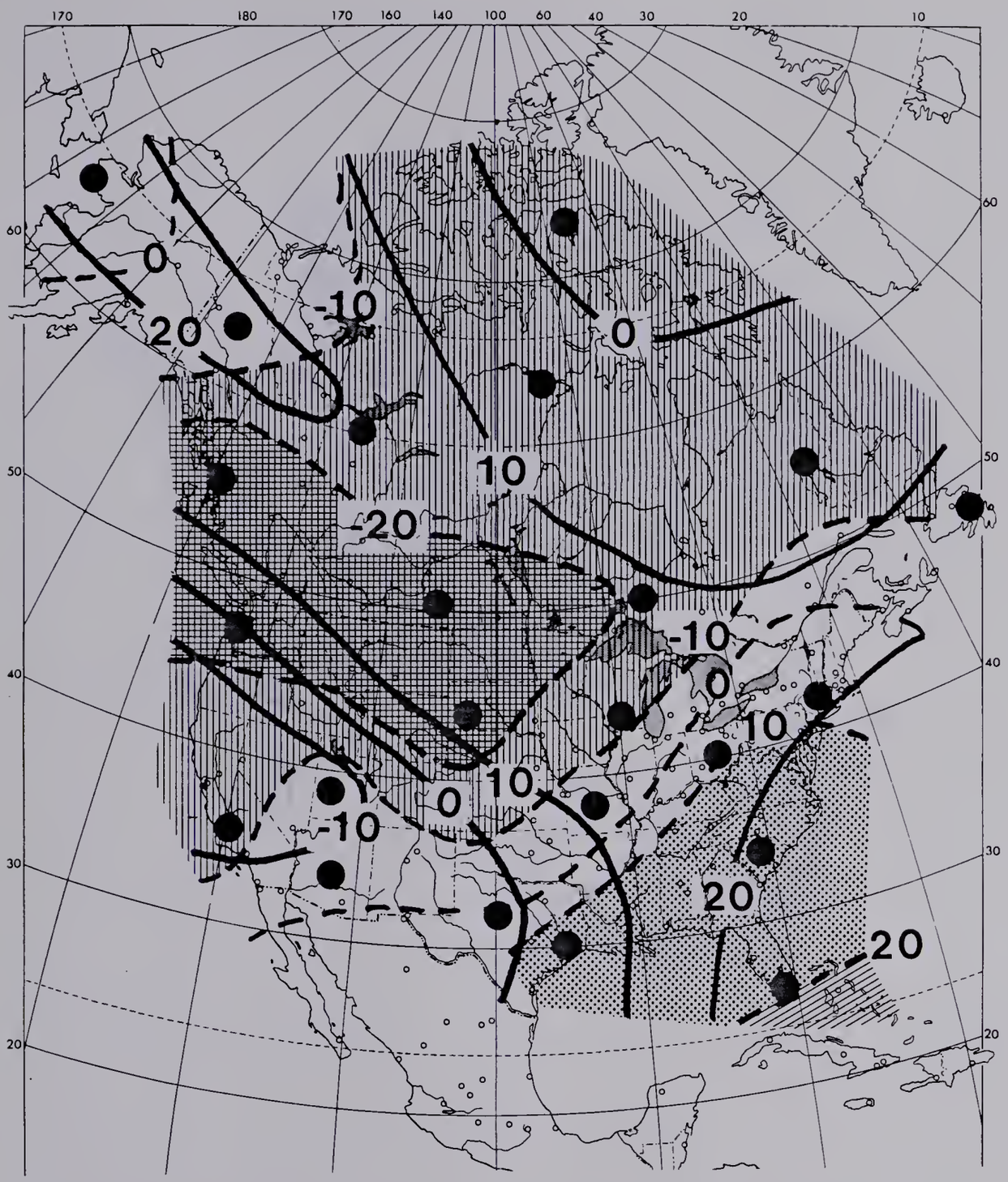

Figure 11. First eigenvector of sea-level pressue and surface temperature. The isolines of pressure are the solid lines, those of temperature are the dashed lines. Regions of maxima or minima in the temperature patterns are identified by stippling or hatching, respectively. (After Kutzbach, 1967)

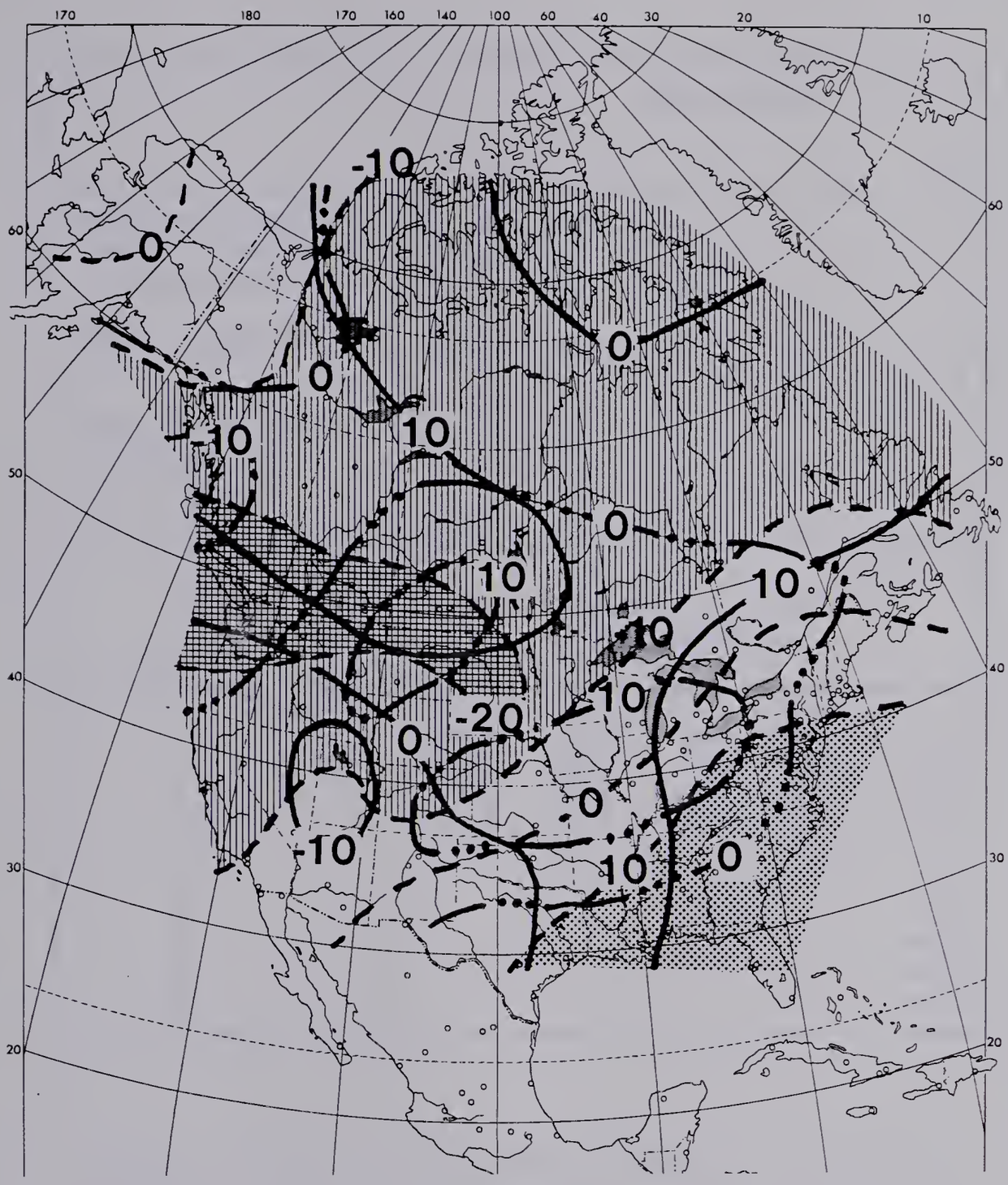

Figure 12. First eigenvector of sea-level pressure, surface temperature and precipitation. Isolines of pressure, temperature and precipitation are indicated by solid lines, dashed lines and dash-dotted lines, respectively. Regions of maxima or minima in temperature are as in Figure 11. (After Kutzbach, 1967)

	Eigenvectors				
	P	T	R	PT	PTR
Number of climatic variables specified at each point J	1	1	1	2	3
Number of points on map K	23	23	23	23	23
Total number of variables M, M=J times K	23	23	23	46	69

k	Cumulative per cent of total variance V_k				
1	36.2	36.4	21.0	28.6	24.3
2	58.7	56.2	34.9	47.8	40.5
3	73.0	70.8	46.8	63.0	53.5
4	79.6	82.6	55.8	72.4	62.0
5	85.6	87.9	63.6	79.4	69.0
6			69.8	83.5	73.0
7			73.5		76.7
8			79.1		80.0
9					83.0
10					85.5

Table 5. Summary of the number of climatic variables and the total number of variables used in various models (top); and the cumulative per cent of total variance V_k explained by the eigenvectors associated with the k largest eigenvalues of their respective correlation matrices. (After Kutzbach, 1967)

variables, the number of eigenvectors required to explain a specified portion of the total variance is inversely related to the degree of intercorrelation between the M variables."

Kutzbach listed four reasons why difficulties might arise in interpreting the patterns: the number of observations was small; the number of points was small; "the distribution of explained variance using only, say, the first five eigenvectors of PTR is not uniform from meteorological variable to meteorological variable or from point to point"; and normalization was perhaps not the optimum weighting scheme. With respect to the last point, Kutzbach repeated his calculations after first weighting the variables so that their average variances were equal, and found that the features of the first several patterns were similar to the previous patterns, though the gradients were greater around points with larger variances.

Kutzbach examined the time coefficients of the PTR eigenvectors and commented that, since only the first four eigenvectors resembled the actual observed normalized departure fields, perhaps the higher order eigenvectors were a result of the orthogonality constraint. He also pointed out that the mean monthly averages are the sum of many weather regimes and hence more than one eigenvector is needed in most cases to represent any one map.

In conclusion, Kutzbach emphasized the interpretability of the eigenvector patterns of the normalized departure fields and suggested their use in "...descriptive or diagnostic studies in which the

interrelationships between fields of several variables are not clearly understood."

2.6 Wallace, J.M. and Dickinson, R.E. (1972)

Wallace and Dickinson extended the application of empirical orthogonal functions to the analysis of the cross-spectra of filtered time series, a procedure they called "complex eigenvector analysis". Their aim was to find "some objective way to define the number of significant wave disturbances present in certain frequency intervals and to separate the total disturbance field into individual wave components".

Since the cross-spectrum contains the power spectrum of each time series as well as their co-spectra and quadrature spectra, they also noted that "...this method should fully exploit the statistical information contained in the cross-spectrum matrix".

After reviewing the usual theory of empirical orthogonal functions, Wallace and Dickinson derived the theory of employing the complex eigenvectors of the cross-spectrum matrix to represent an augmented time series which was derived from the original data. The augmented time series, which was complex, was used instead of the original data in order to generate time coefficients which were real rather than complex, and hence facilitate further cross-spectrum analysis of the time series. The authors pointed out that the method used to derive the theory applied only at one frequency and so, if other frequencies were to be studied, different eigenvectors

would have to be generated for each frequency, or each frequency interval. Also mentioned is the need for some form of normalization if time series of different variables were to be analyzed, though the work of Wallace (1972) seems to indicate that the results obtained are independent of the normalization scheme used.

Wallace and Dickinson also examined the interpretation of the wave structures represented by the complex eigenvectors. Noting that some modes can be shown to be statistically significant and the rest rejected as noise, the authors point out that "statistical significance does not necessarily guarantee physical significance" and that the results have to be compared with results from "synoptic and/or dynamical modeling studies". If one wave structure were present, then the first mode should represent all the information about the wave, with the higher modes representing noise, but if two waves were present then complex eigenvector analysis may or may not be able to separate them, and ordinary eigenvector analysis would be incapable of detecting one of the waves under certain conditions. In view of the fact that the results of complex eigenvector analysis may or may not be interpretable, Wallace and Dickinson looked at two situations.

First they asked if it were possible to determine if any wave structures at all were present in a set of data, in a given frequency band, by the use of complex eigenvector analysis. If the first eigenvalue turned out much larger than any of the others, then, Wallace and Dickinson reasoned, one wave structure must surely be present. Alternatively, if the first few eigenvalues were

larger than the following ones, several wave structures may be present and, in order to separate them, more input parameters would likely be required.

The other situation commented on was the testing of hypotheses concerning wave structures. Wallace and Dickinson infer that it should be possible to select input parameters which will produce information about the composition of the waves and hence produce a means of testing hypotheses.

The authors remark that the selection of input parameters is important, since the wrong choice may yield wave structures which are not highly orthogonal and therefore indistinguishable. Orthogonal wave structures were explained by the following example.

Suppose two superimposed waves, one with a long wavelength, the other with a short wavelength are moving eastward, and the area of data points is such that it encompasses at least one wavelength of the shorter wave, yet small enough so that changes in the longer wave are simultaneous at all stations. If the input parameters consisted of observations of stations on a north-south line or of observations of one station at different levels, the two wave structures would be indistinguishable and hence "the wave structures would have no orthogonality with respect to this particular set of input parameters".

Wallace and Dickinson comment that if an optimum combination of input parameters cannot distinguish two or more wave structures by means of complex eigenvector analysis, then any other method would also fail.

Some of the problems dealt with by real eigenvector analysis could possibly be better handled by the complex method, the authors conclude, since more information is provided by complex eigenvector analysis and the information can be limited to one frequency interval.

2.7 Wallace, J.M. (1972)

As an example of the possible application of complex eigenvector analysis, Wallace analyzed data from 12 stations in the tropical Pacific region for the period July to October, 1967. Using conventional cross-spectrum analysis the author previously found three distinct disturbances of periods 4 to 5 days: "mixed Rossby-gravity waves", "synoptic-scale, westward propagating disturbances associated with the intertropical convergence zone", and "synoptic-scale, westward propagating disturbances of subtropical latitudes".

The available input data consisted of daily values of the zonal wind component, the meridional wind component, the temperature at the surface and at 22 pressure levels from 950 to 70 millibars, the satellite viewed cloud brightness, the vertically averaged relative humidity, the total cloud cover, the opaque cloudiness and the precipitation. The data were normalized to unit variance in the frequency interval of interest.

Without going into the specifics of Wallace's results, it is sufficient to mention that the author's conclusions were consistent with his previous work using cross-spectrum analysis, as well as

with the work of other researchers using other methods. Moreover, some differences between early and recent studies were clarified and some uncertainties in the author's past results were resolved.

CHAPTER 3

APPLICATIONS

3.1 Introduction

An analysis of 4 sets of data was performed in order to demonstrate the use of empirical orthogonal functions in a Canadian context and to draw some conclusions from the results. The data include a simple hypothetical wave pattern, ten years of monthly mean temperatures from Western Canada, ten years of 500-mb contour-height summer data from Western Canada and the North-Western U.S. and ten years of precipitation data for the same area and period as the 500-mb heights.

3.2 A Simple Wave Pattern

The wave pattern of Figure 13, representing a pressure field moving from left to right was analyzed for the nine indicated stations (black dots) and two cycles (24 time units). The input values were read from the figure rather than calculated.

The first two time functions are plotted in Figure 15 with the corresponding two eigenvectors shown in Figure 14. These two eigenvectors explained over 96 per cent of the variance with the other 4 per cent being explained by inaccurate readings of values from the initial data.

At least two conclusions can be drawn from this exercise. One is that if two time functions correlate fairly highly at some lag

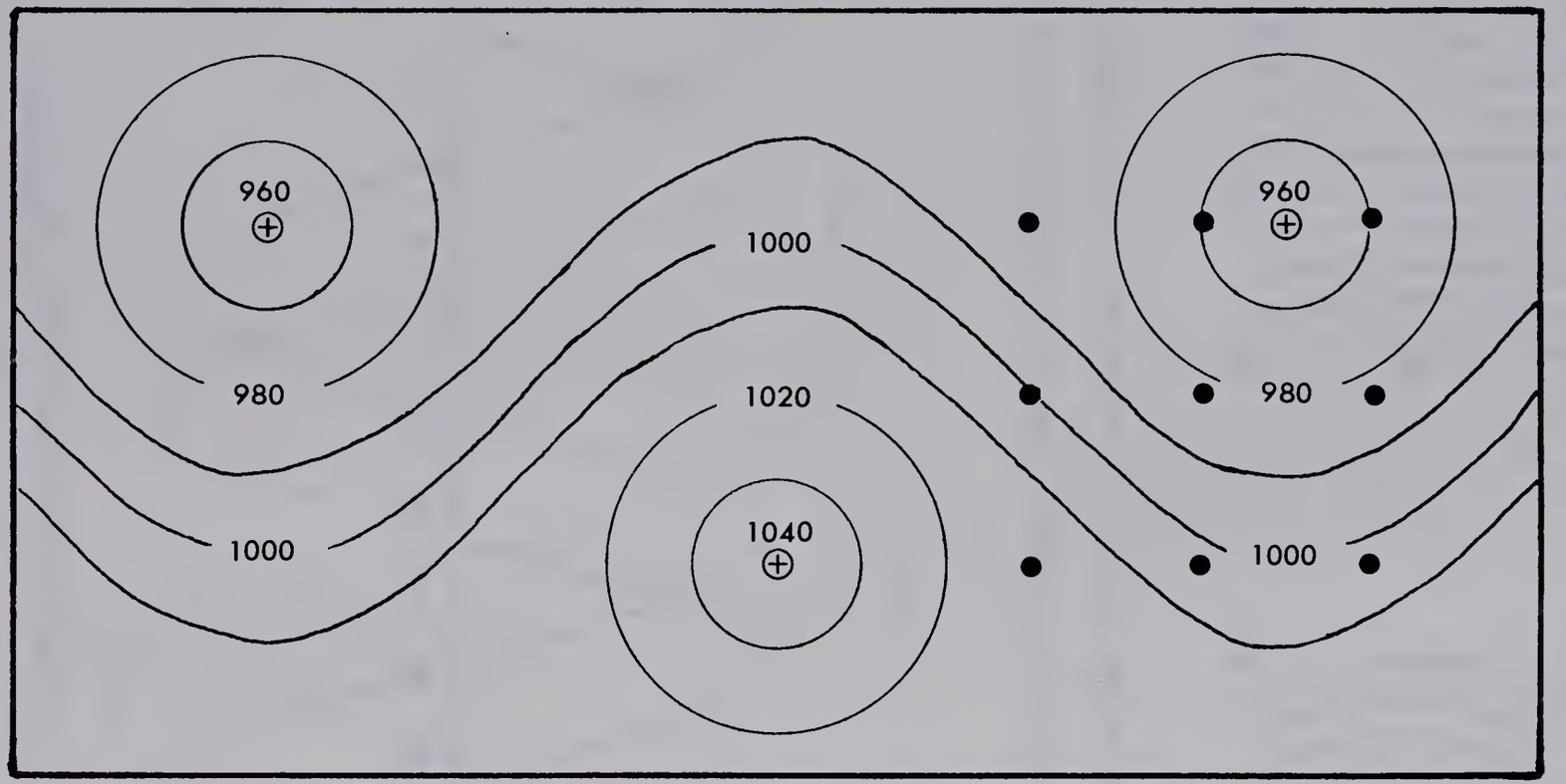

Figure 13. Hypothetical pressure field moving left to right with a period of 12 time units. Dots indicate stations.

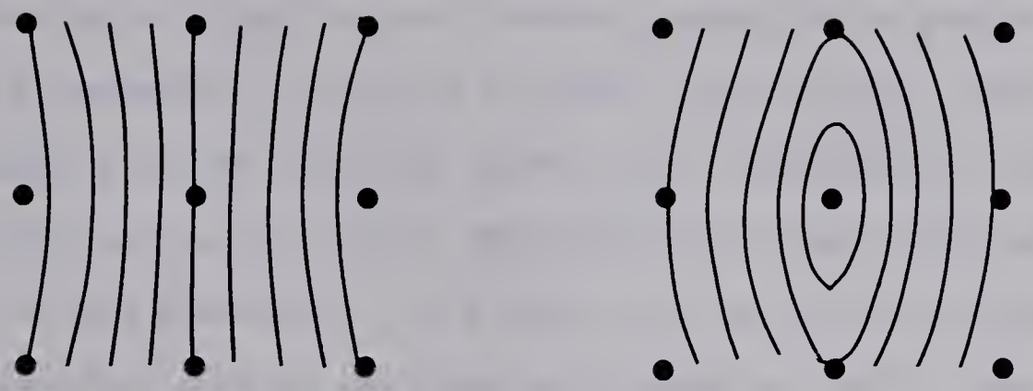

Figure 14. Eigenvector No. 1 (a), and eigenvector No. 2 (b) derived from the wave pattern of Figure 13.

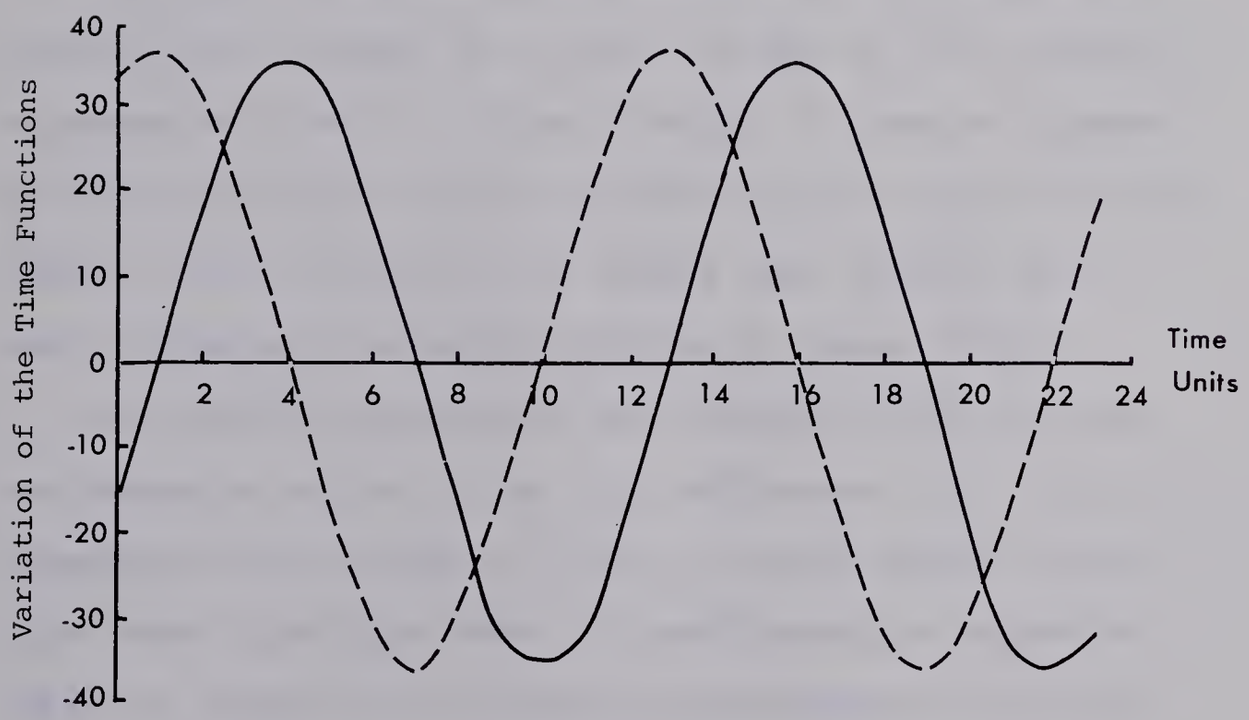

Figure 15. Time functions of the eigenvectors in Figure 14. Time function No.1 and time function No. 2 are the solid and dashed lines, respectively.

then there is a possibility that some type of wave phenomenon is being observed. Such a correlation was attempted with some of the later data but with poor results. Another possibility is that if the two eigenvectors of Figure 14 represent Figure 13 then perhaps other simple patterns like those formed by the eigenvectors in Figure 14 may exist and can aid in the identification of some predominant yet not obvious phenomenon. For example, if a set of data produced some eigenvector pattern, say a series of parabolas, and if a completely different set of data produced the same parabolas, then perhaps the same sort of structure may be present and observed in both sets of data.

3.3 Ten Years of Monthly Mean Temperatures

Mean monthly temperatures were extracted for the twenty-five stations shown in Figure 16 for the years 1963 to 1972, inclusive. The variances, Figure 17, and means, Figure 18, show the expected spatial distributions; variances lowest along the coasts, then increasing inland, while the means exhibit their customary north-south gradient with an easterly tilt due to oceanic effects.

The twenty-five eigenvectors were calculated after the means were removed as well as all the time coefficients. The first five eigenvectors are displayed in Figures 19 through 23 and the first three years of the first three time coefficients are graphed in Figure 24. Figures 25 to 28 show the periodograms of the first four time functions. Also, Table 6 shows the contribution of each eigenvector to the total variance.

Figure 16. The 25 stations used in analyzing 10 yrs. of monthly mean temperatures.

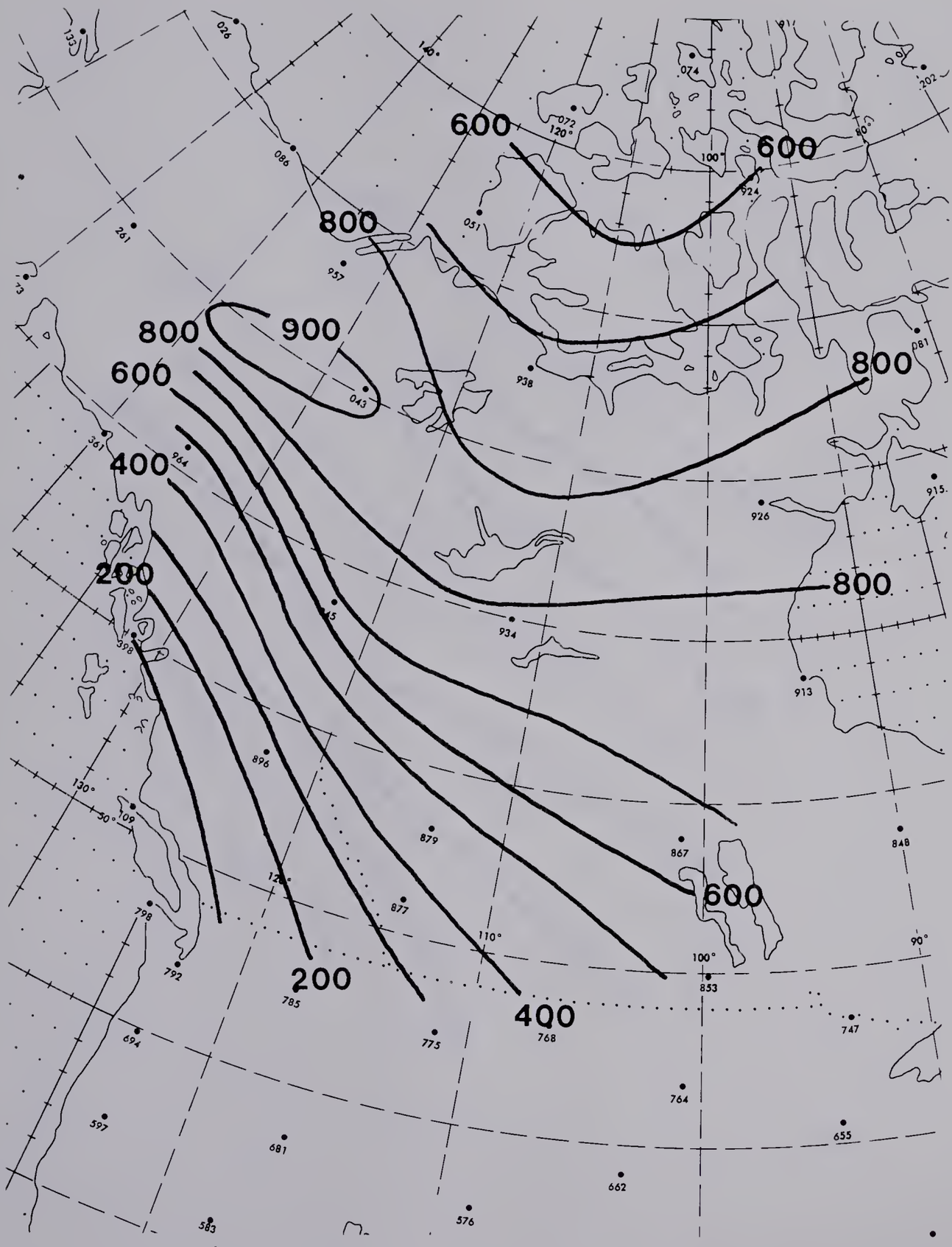

Figure 17. Variances of monthly mean temperatures in $(^{o}F)^{2}$.

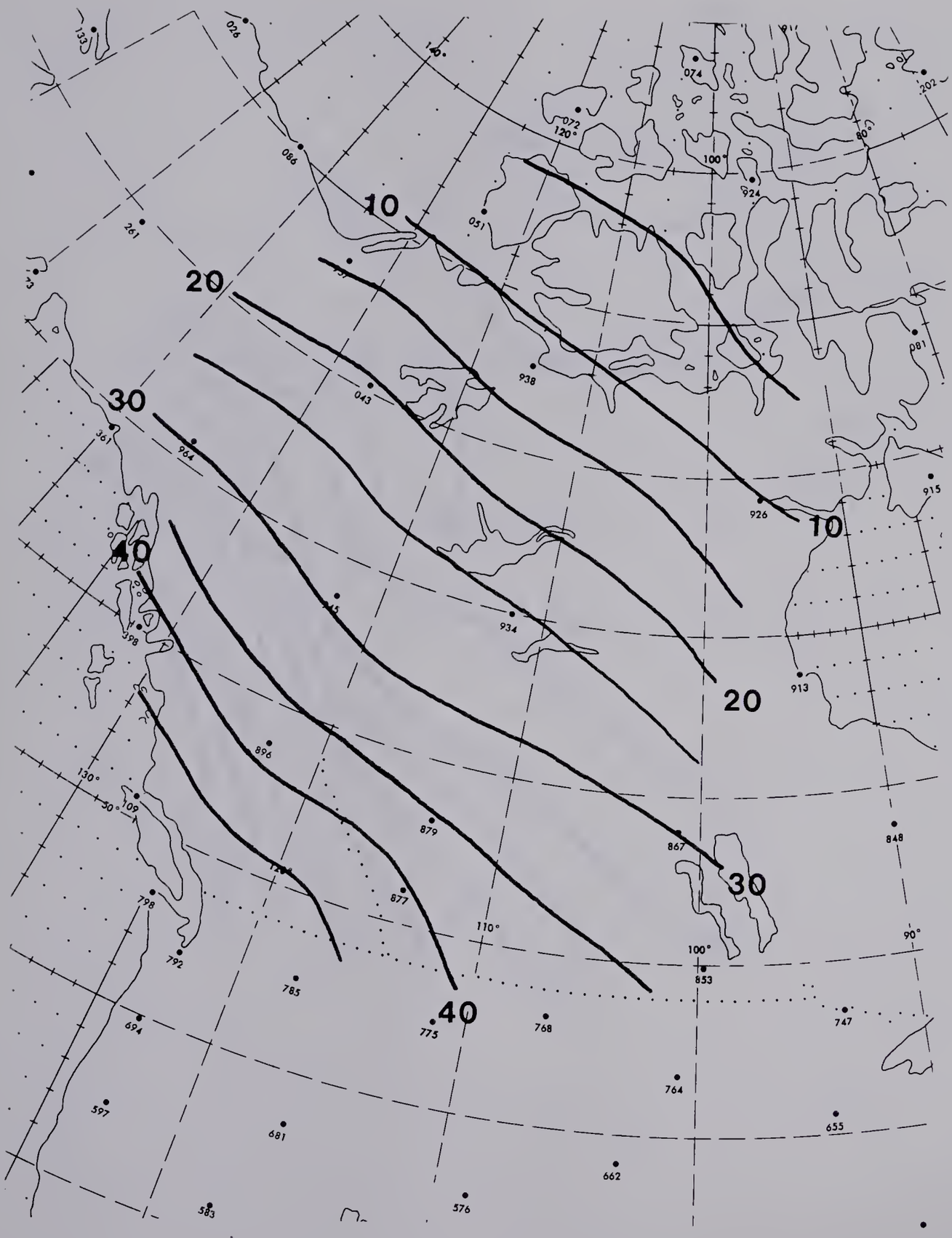

Figure 18. Means of monthly mean temperatures in ^{o}F.

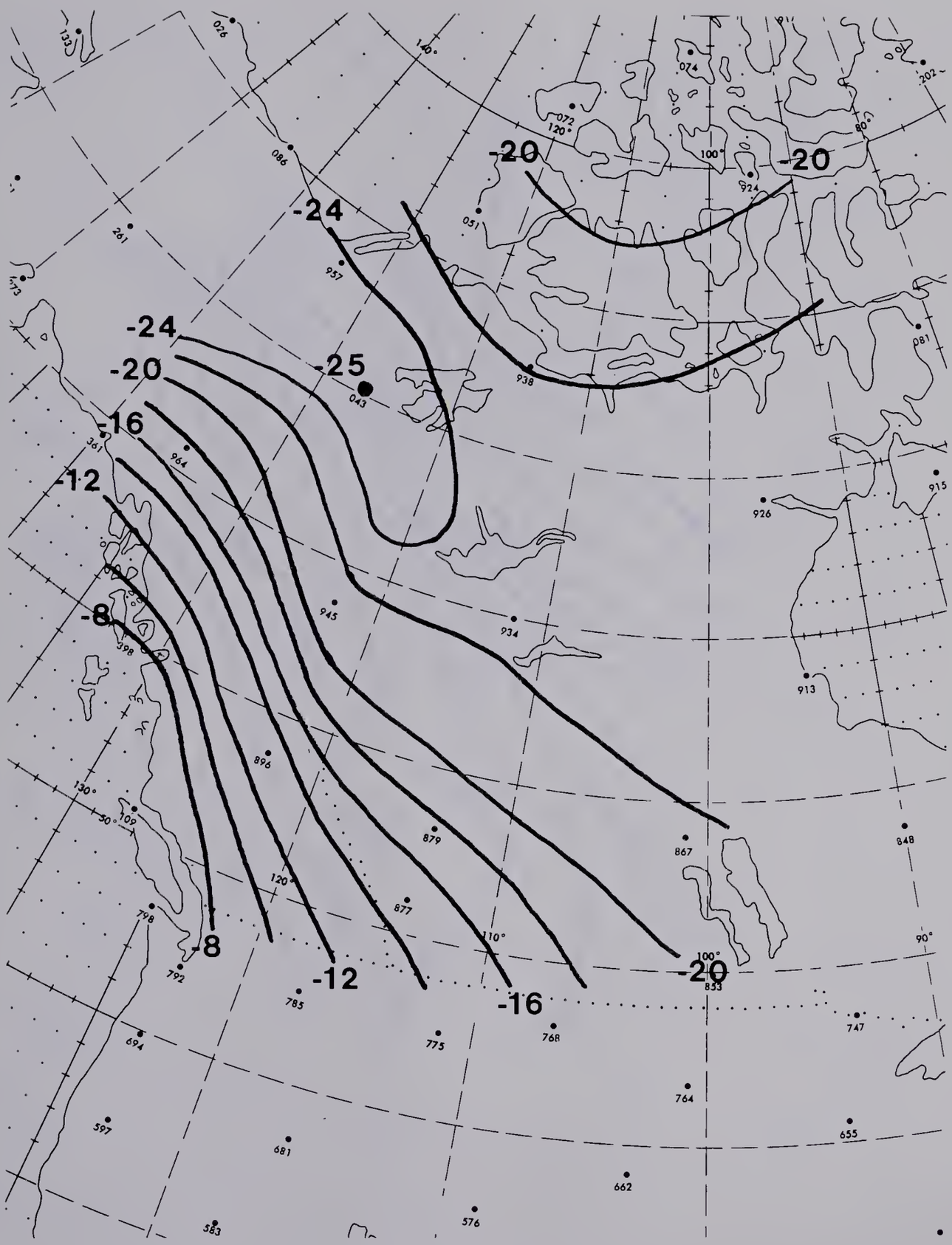

Figure 19. Eigenvector No. 1 of the temperature data. ($x10^2$)

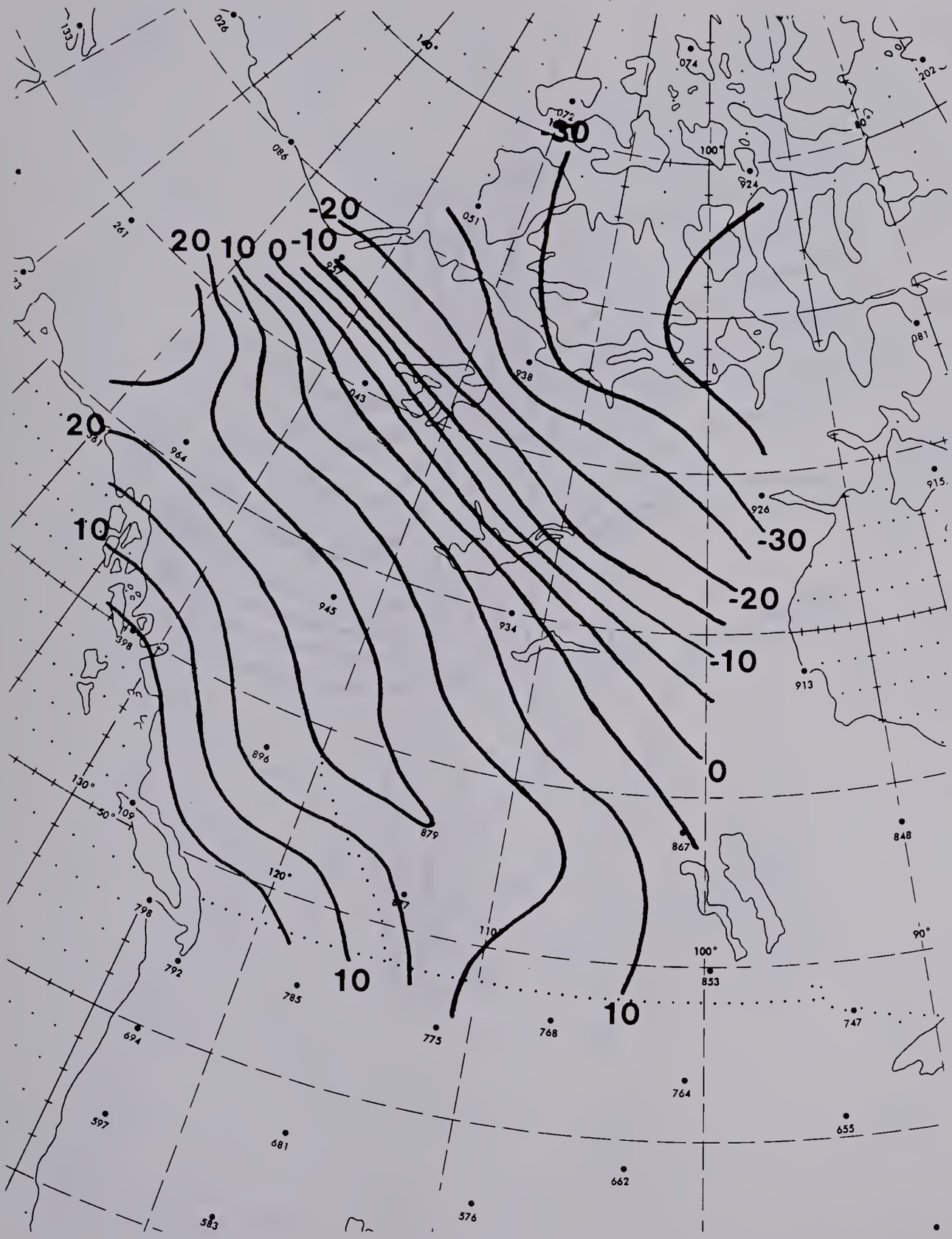

Figure 20. Eigenvector No. 2 of the temperature data. ($x10^2$)

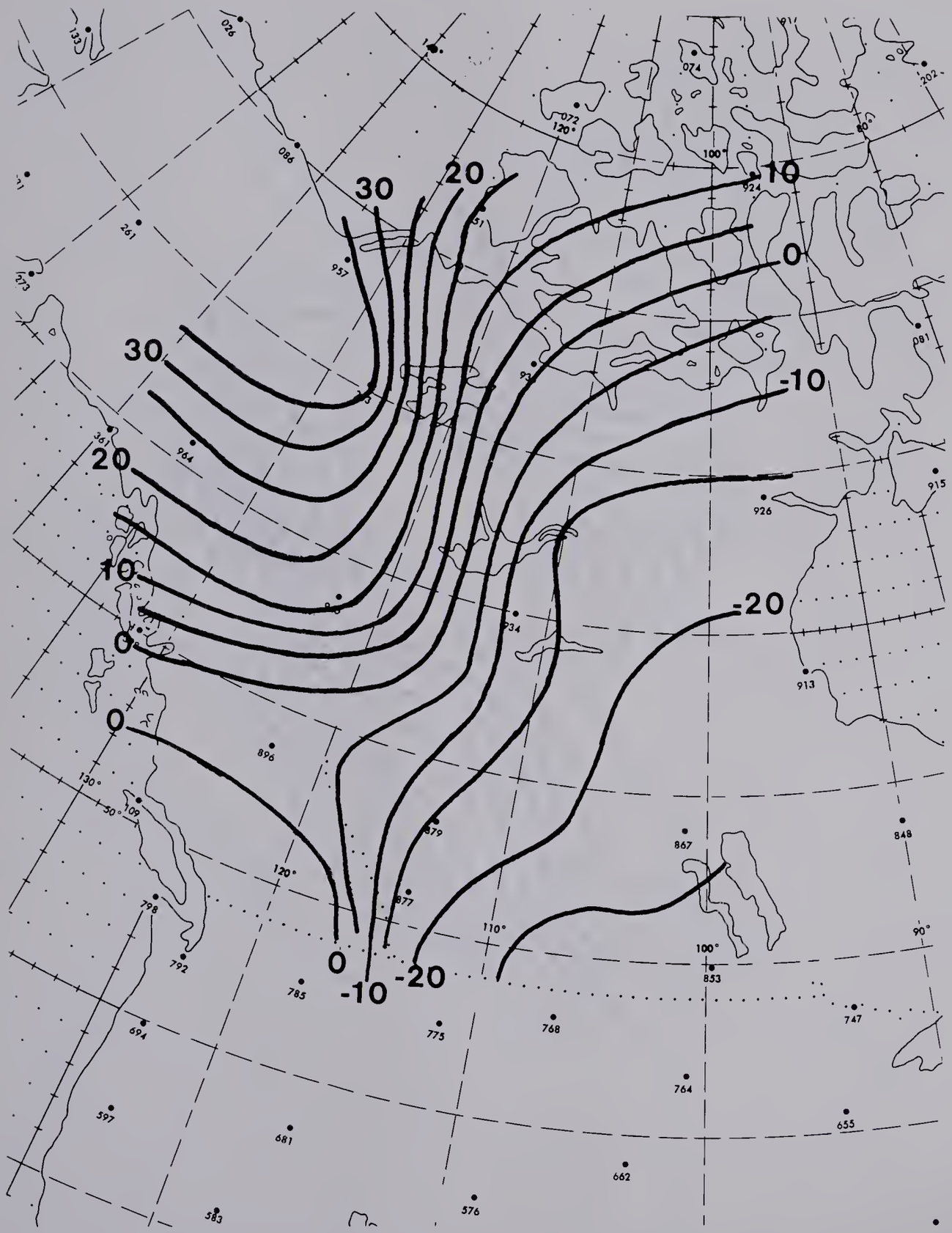

Figure 21. Eigenvector No. 3 of the temperature data. ($x10^2$)

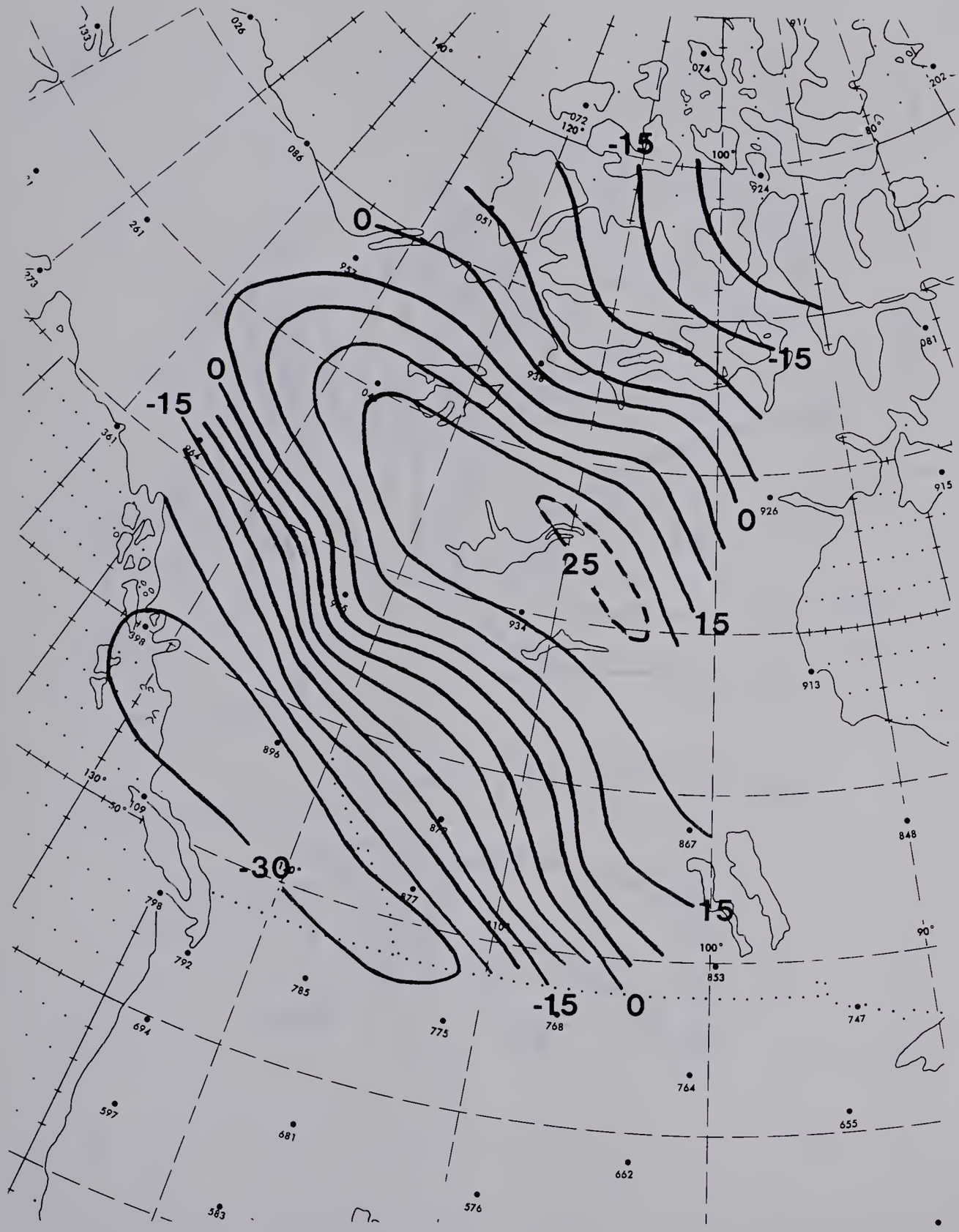

Figure 22. Eigenvector No. 4 of the temperature data. ($x10^2$)

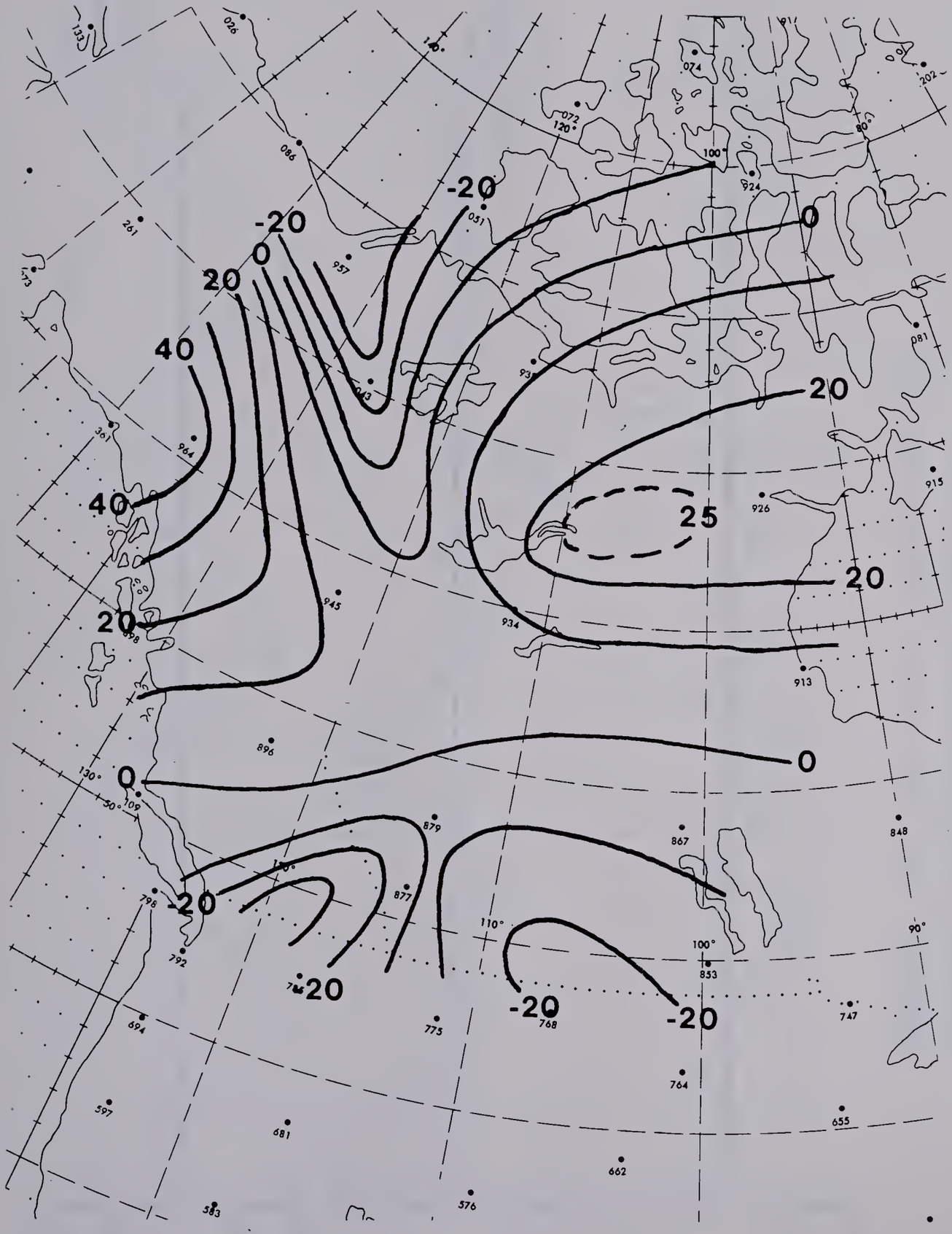

Figure 23. Eigenvector No. 5 of the temperature data. ($x10^2$)

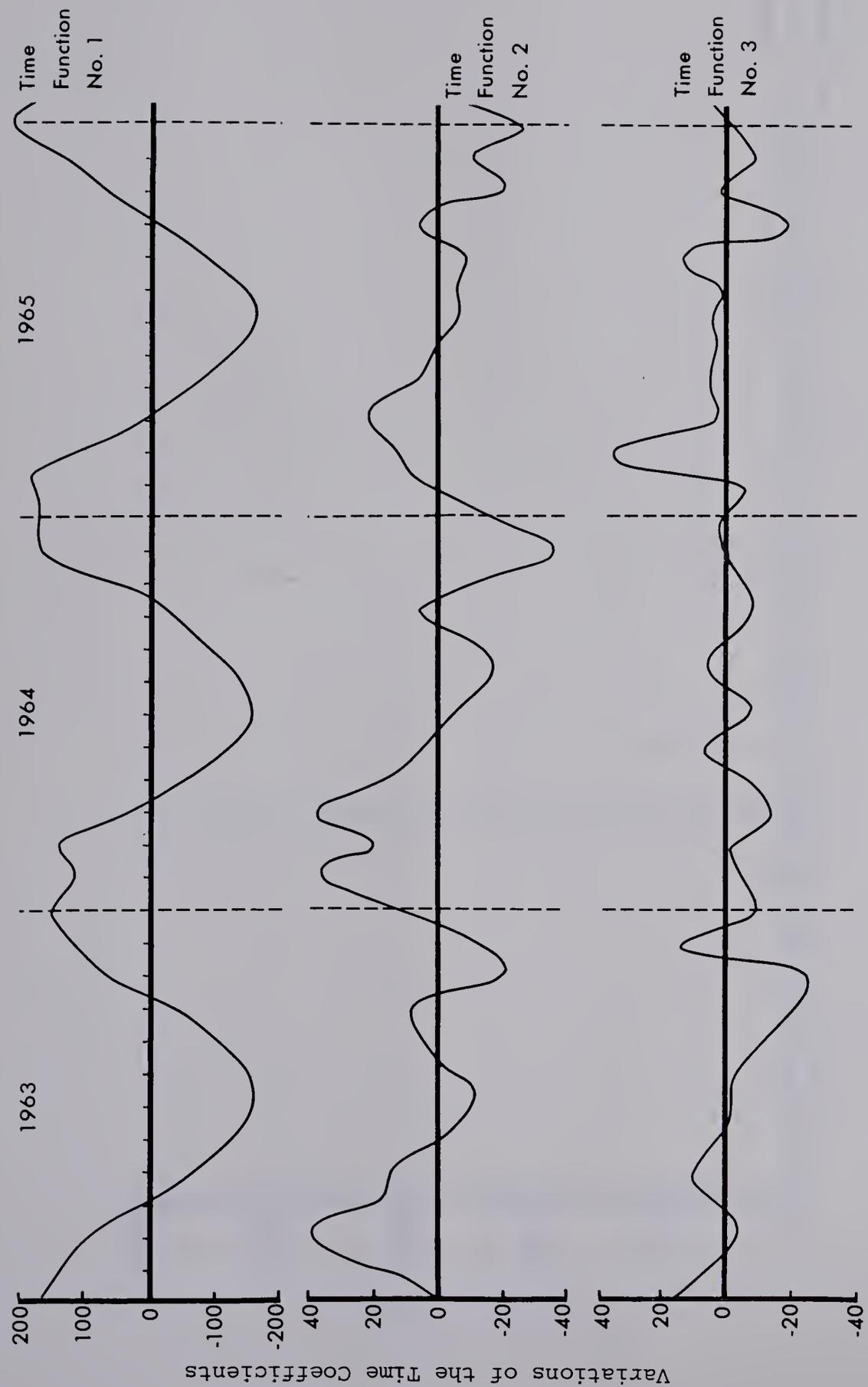

Figure 24. First three years of the first three time coefficients of the monthly mean temperatures.

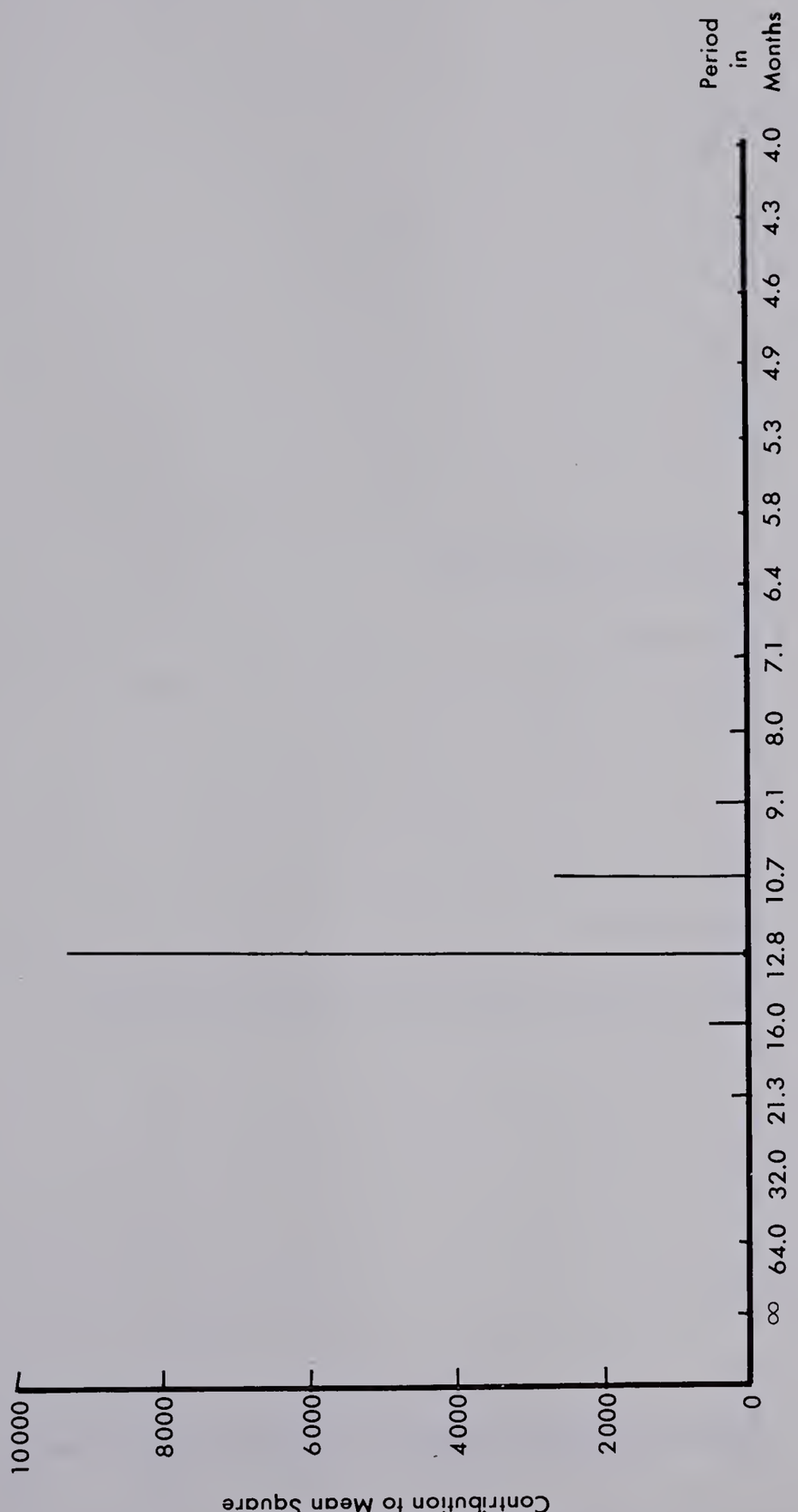

Figure 25. Periodogram of time function No.1 of the temperature data.

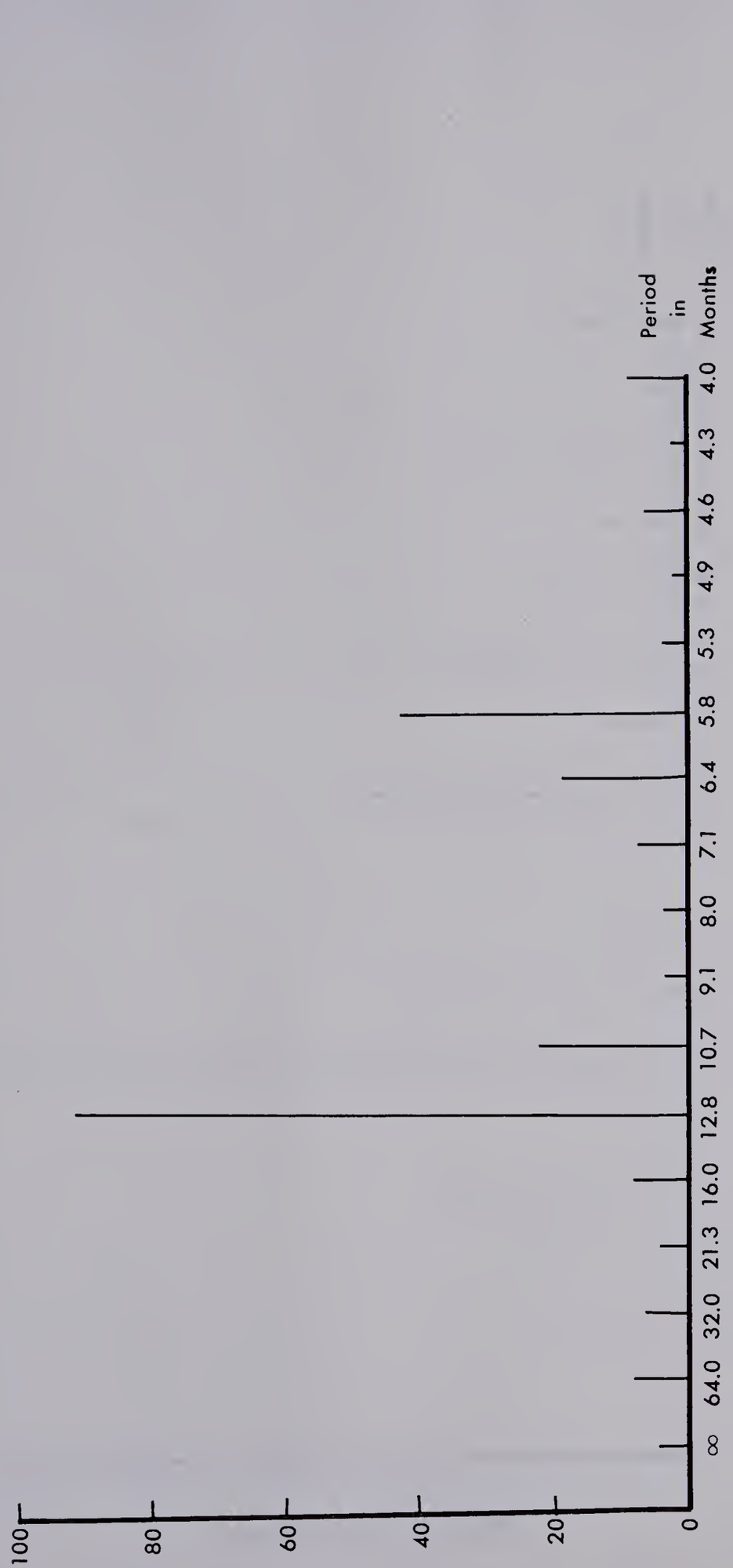

Figure 26. Periodogram of time function No. 2 of the temperature data.

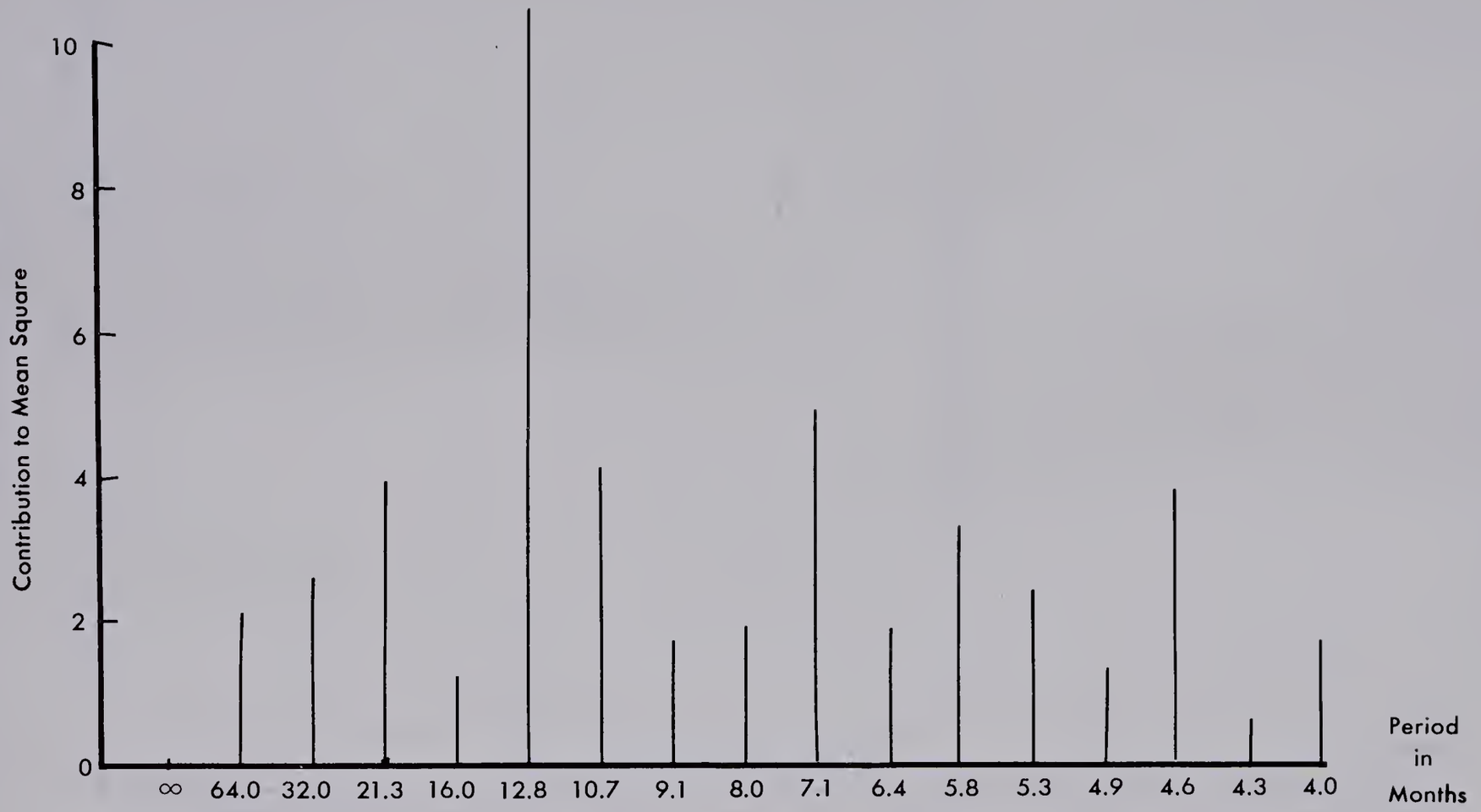

Figure 27. Periodogram of time function No. 3 of the temperature data.

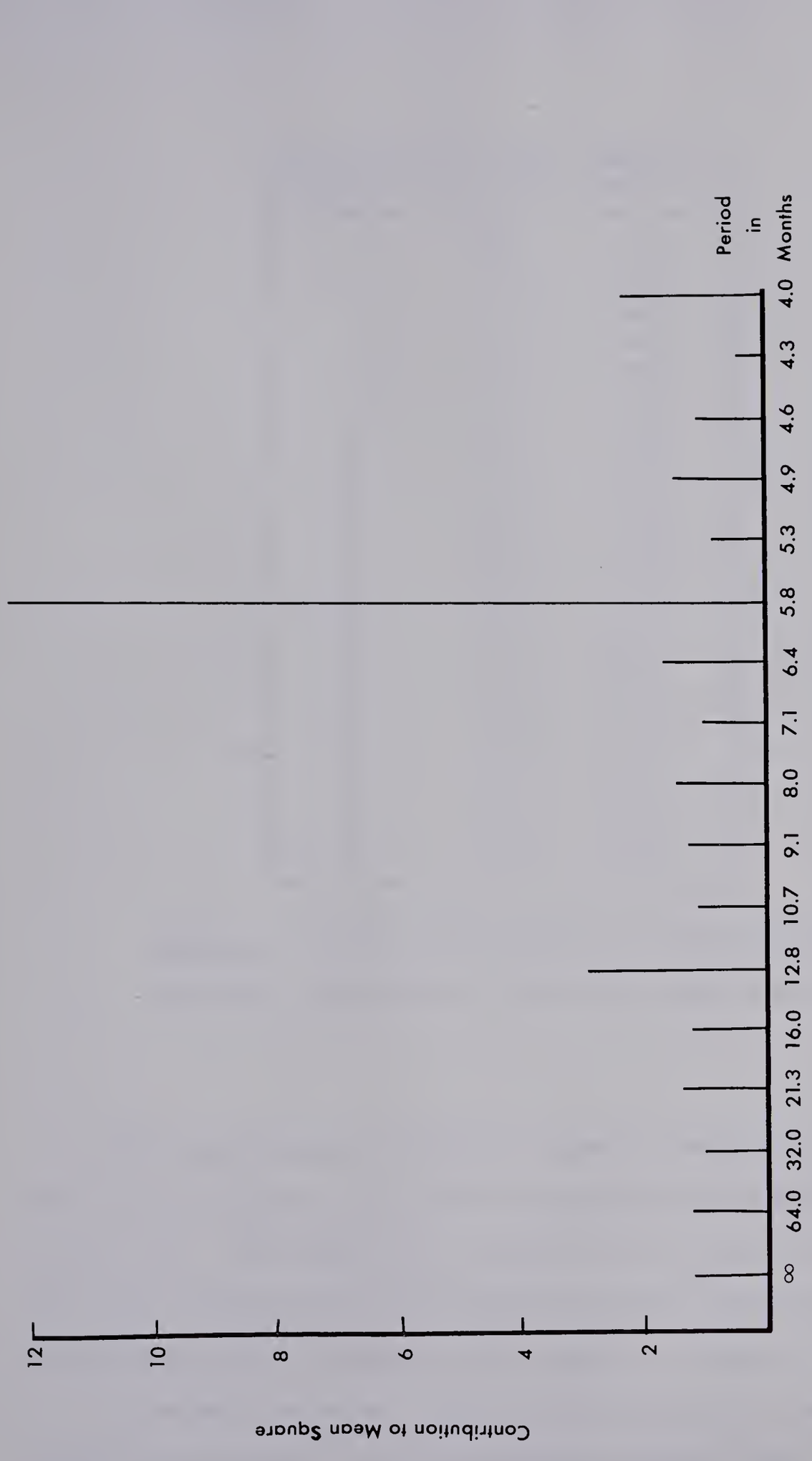

Figure 28. Periodogram of time function No. 4 of the temperature data.

Eigenvector Number	Per Cent Variance	Cumulative Variance
1	96.4	96.4
2	2.0	98.4
3	0.6	99.0
4	0.3	99.3
5	0.2	99.5
6	0.1	99.6
7	0.1	99.7
8	0.1	99.8
9	0.1	99.8
10	0.0	99.8
11	0.0	99.9
12	0.0	99.9
13	0.0	99.9
14	0.0	99.9
15	0.0	99.9
16	0.0	100.0
17	0.0	100.0
18	0.0	100.0
19	0.0	100.0
20	0.0	100.0
21	0.0	100.0
22	0.0	100.0
23	0.0	100.0
24	0.0	100.0
25	0.0	100.0

Table 6. Variances and cumulative variances explained by successive eigenvectors of the monthly mean temperatures.

The first eigenvector, Figure 19, closely resembles the variance map. Such a similarity was pointed out previously by Craddock and Flood (1969). This eigenvector accounted for over 96 per cent of the total variance and its associated time function exhibited a strong yearly cycle as shown by the periodogram of Figure 25.

The second eigenvector, Figure 20, contributed only 2 per cent to the total variance, but its time function exhibited some

interesting properties. First, the periodogram of the time function (Figure 26) indicates a strong yearly variation coupled with a weaker six-month variation which is unusual and not easily explained. Second, the eigenvector pattern has a zero line running from the north-west to the south-east which means that for part of the year this eigenvector raises/lowers temperatures towards the north-east and lowers/raises them towards the south-west by about 10°F.

Eigenvectors No. 3 and 4 do not seem to show any readily recognizable features although the time function of eigenvector No. 4 shows a strong semi-annual variation. Eigenvector No. 5, although explaining only 0.2 per cent of the variance, bears some discussion since it could possibly be representing a real phenomenon. Although the network of stations is not dense enough to allow any definite conclusions, it appears that the trough at the top of the pattern almost exactly follows the MacKenzie River. Considering the signs of the eigenvectors, the temperatures along the river are a bit higher when the temperatures to the east and west are lower, and are lower when the temperatures to the east and west are higher, anywhere from less than a degree to about 10°F. If, indeed, this is the case, then the problem of what is to be considered as noise and what is to be be considered as "real" must be re-examined, since even with 99.5 per cent of the variance explained, real processes might be rejected.

The eigenvectors and time functions were now recombined in order to examine the errors that would result if not all the functions were used. When only the first eigenvector was used 113 errors resulted that were greater than 10°F, 20 that were greater than 15°F

and one a maximum of 22°F. Using two components, explaining 98.4 per cent of the variance, only 19 errors greater than 10°F were produced, with one error over 15°F. By the tenth component, explaining 99.8 per cent of the variance, there were 38 errors over 2.5°F, none being more than 4°F.

If these errors are assumed to be "noise", possibly caused by faulty data handling and tabulation, rather than actual anomalies, then orthogonal functions are ideal for storing these sorts of climatological data. Instead of storing ten years x twelve months/year x 25 values/month = 3000 values, only the eigenvector matrix of 10 x 25 = 250 values plus the first ten time coefficients of 1200 values need be stored, resulting in a saving of space of over 50 per cent. As more data are gathered, only the time functions need be stored as the eigenvectors are not expected to change, so that the percentage of storage space saved increases with time.

The critical question still remains: what if the next eigenvector represents a real phenomenon? If data storage by use of eigenvectors were to be adopted then any data assumed to be noise, correctly or incorrectly, would be irretrievably lost.

3.4 Analysis of 500-mb Heights during the Summer Months

For this study nineteen stations (shown in Figure 29) were chosen in western Canada and the north-western United States. The stations were far enough apart so that there would be no redundancy

Figure 29. Stations used in the eigenanalysis of 500-mb heights.

of data, yet close enough together to possibly permit study of smaller-scale effects. The 500-mb heights were obtained from the 0000Z and 1200Z radiosonde reports for each day from May 1 to September 30, for the year 1963 to 1972, inclusive

Figure 30 shows the mean 500-mb field, with the gradient being in the expected north-south direction. The variances are presented in Figure 31. An interesting note is that the minimum variance occurs to the west of the Rocky Mountains while the maxima appear to be in the Gulf of Alaska and to the north-west of Manitoba, perhaps over Hudson's Bay. The isolines of variance shown by Craddock and Flood (1969) run almost perpendicular to those of Figure 31. This may be caused by either the difference in data (Craddock and Flood used 500-mb heights for the full year) or by the coarse grid of Craddock and Flood being unable to pick out what may be an anomaly.

The means were removed and the data were analyzed as before. Table 7 shows the contribution of each eigenvector to the total variance.

Figure 32 illustrates eigenvector No. 1 and shows a positive or negative addition to the mean field, depending on the sign of the time function, centered in southern British Columbia. Its associated time function is plotted in Figure 33 for two years; it appears that it may be part of a yearly cycle. Eigenvector No. 1 is peculiar for two reasons: it shows little resemblance to the variance field of Figure 31 and it shows little resemblance to

Eigenvector Number	Per Cent Variance	Cumulative Variance
1	57.4	57.4
2	16.3	73.6
3	12.2	85.8
4	5.5	91.3
5	2.7	93.9
6	2.1	96.0
7	1.1	97.1
8	0.7	97.8
9	0.5	98.3
10	0.4	98.6
11	0.3	98.9
12	0.3	99.2
13	0.2	99.4
14	0.2	99.5
15	0.1	99.7
16	0.1	99.8
17	0.1	99.9
18	0.1	99.9
19	0.1	100.0

Table 7. Variances and cumulative variances explained by successive eigenvectors of the 500-mb heights.

eigenvector No. 1 of Craddock and Flood (see Figure 9).

Eigenvector No. 2 is displayed in Figure 34. It shows a strong north-west to south-east gradient with a larger positive/negative contribution in the mid-northern United States. Eigenvector No. 3, Figure 35, shows an almost perpendicular gradient to that of eigenvector No. 2 with a positive/negative addition south-west of Vancouver Island, and a negative/positive addition over northern Manitoba. Eigenvector Nos. 4, 5, 6 and 7 are presented in Figures 36 to 39, inclusive.

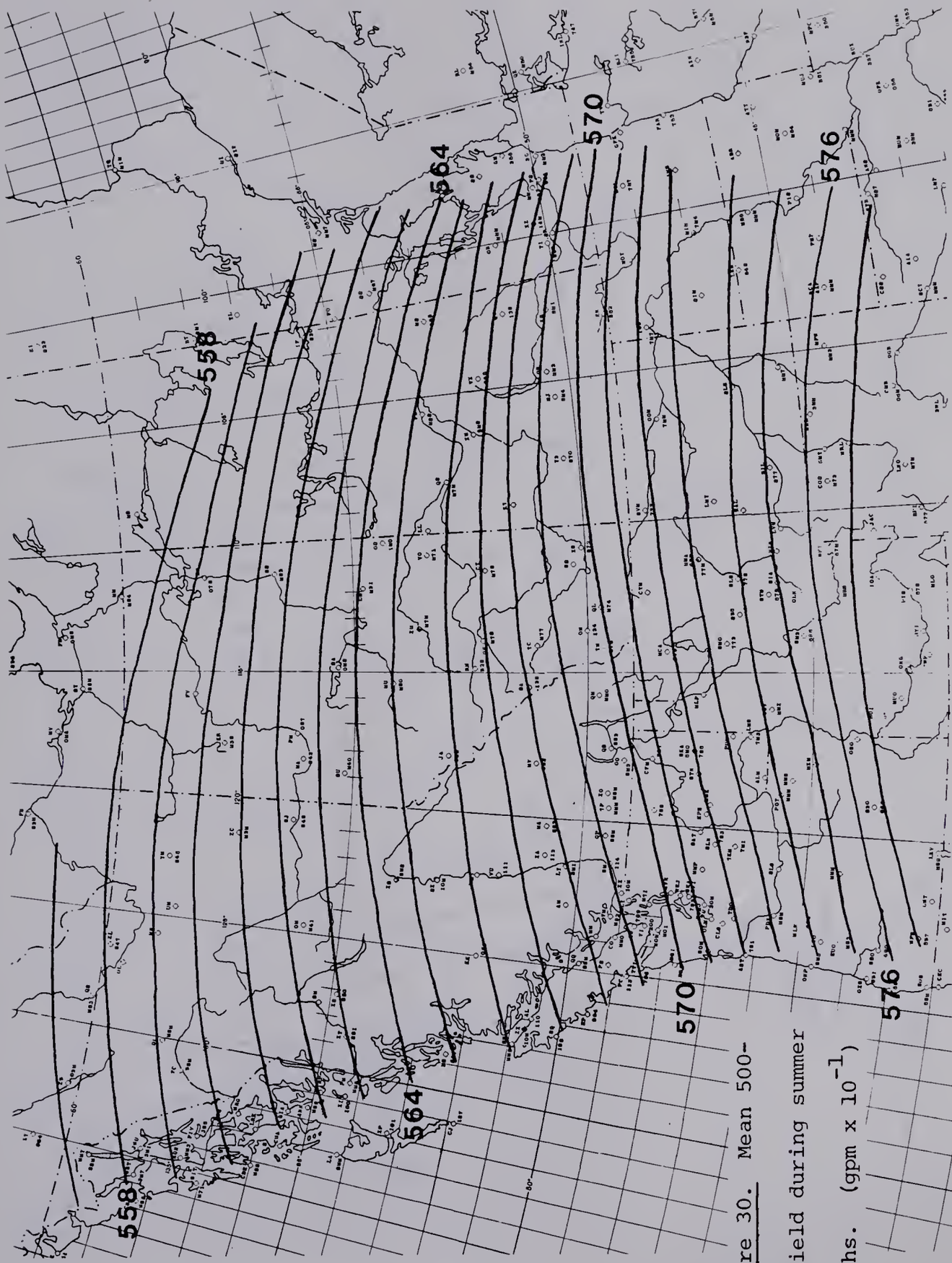

Figure 30. Mean 500-mb field during summer months. (gpm x 10^{-1})

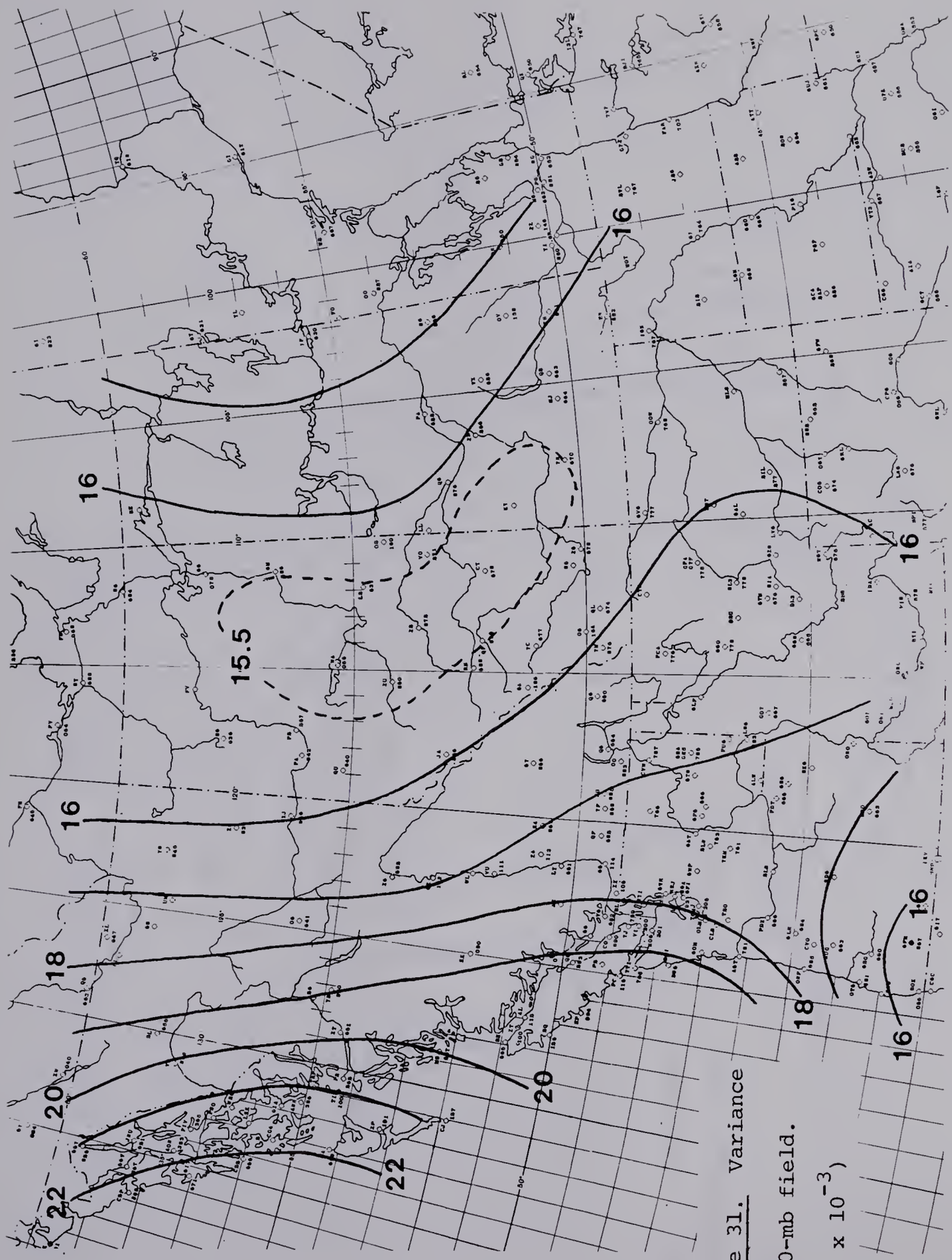

Figure 31. Variance of 500-mb field. ($gpm^2 \times 10^{-3}$)

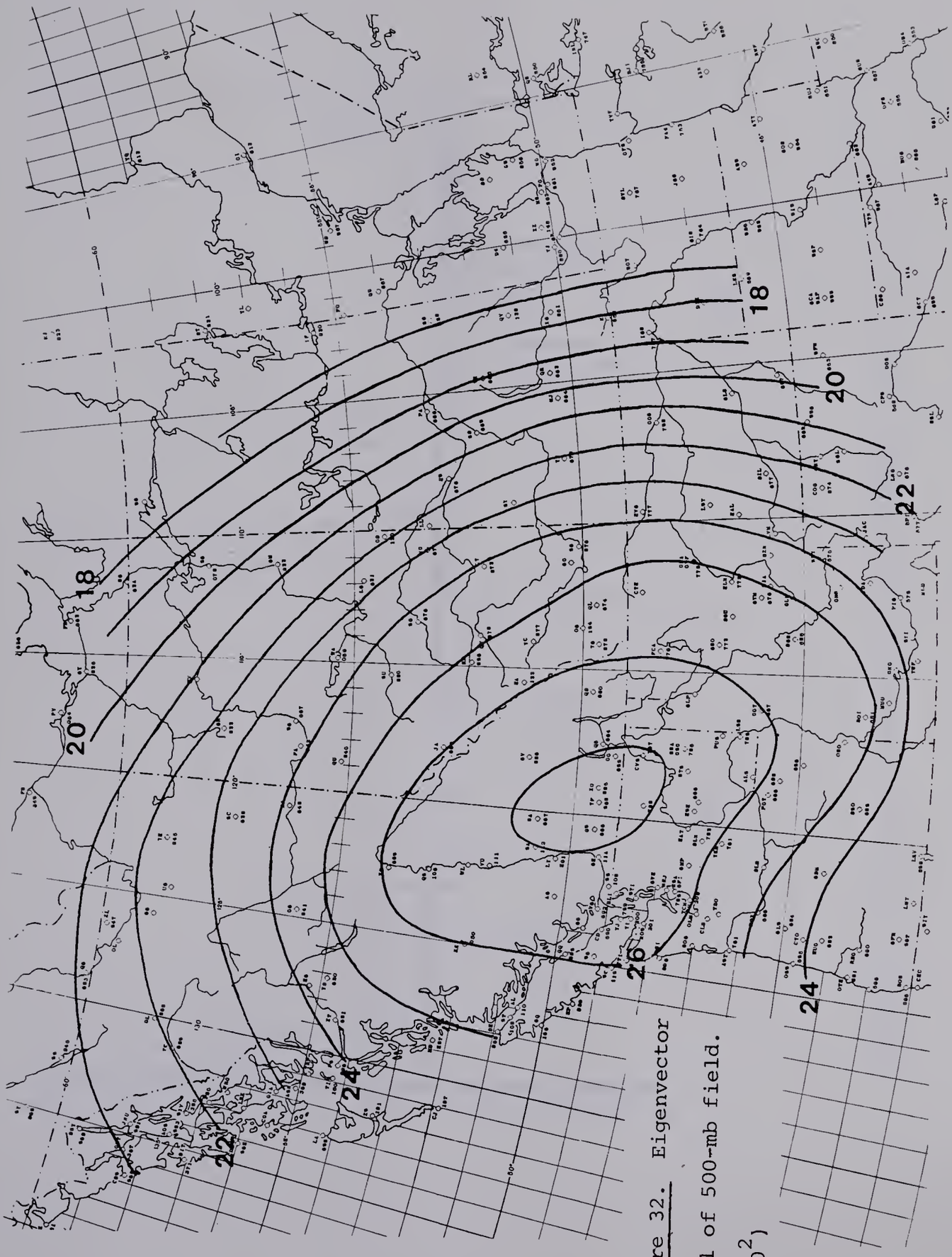

Figure 32. Eigenvector No. 1 of 500-mb field. (x 10^2)

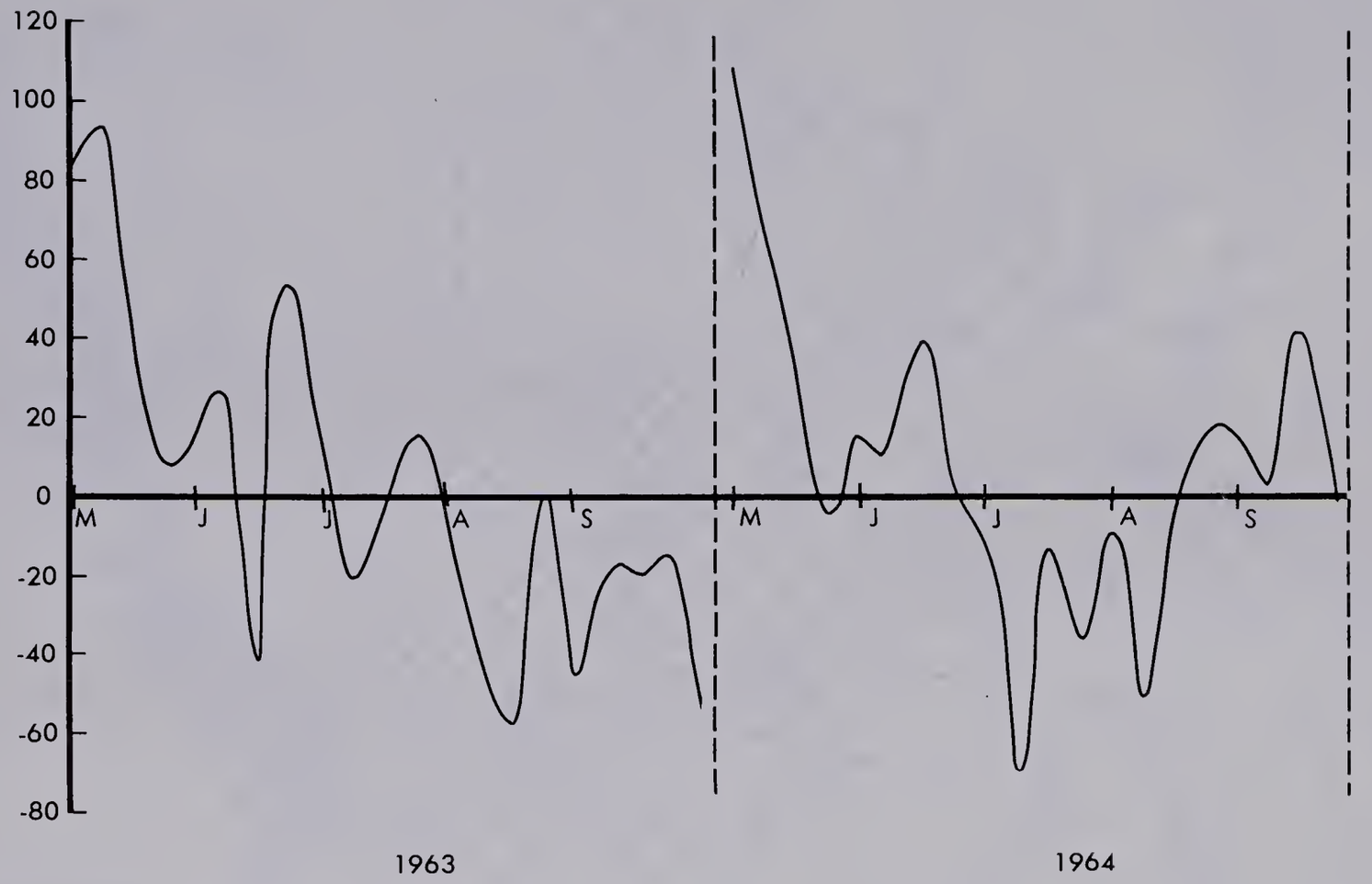

Figure 33. First time function of 500-mb heights, shown for 2 years.

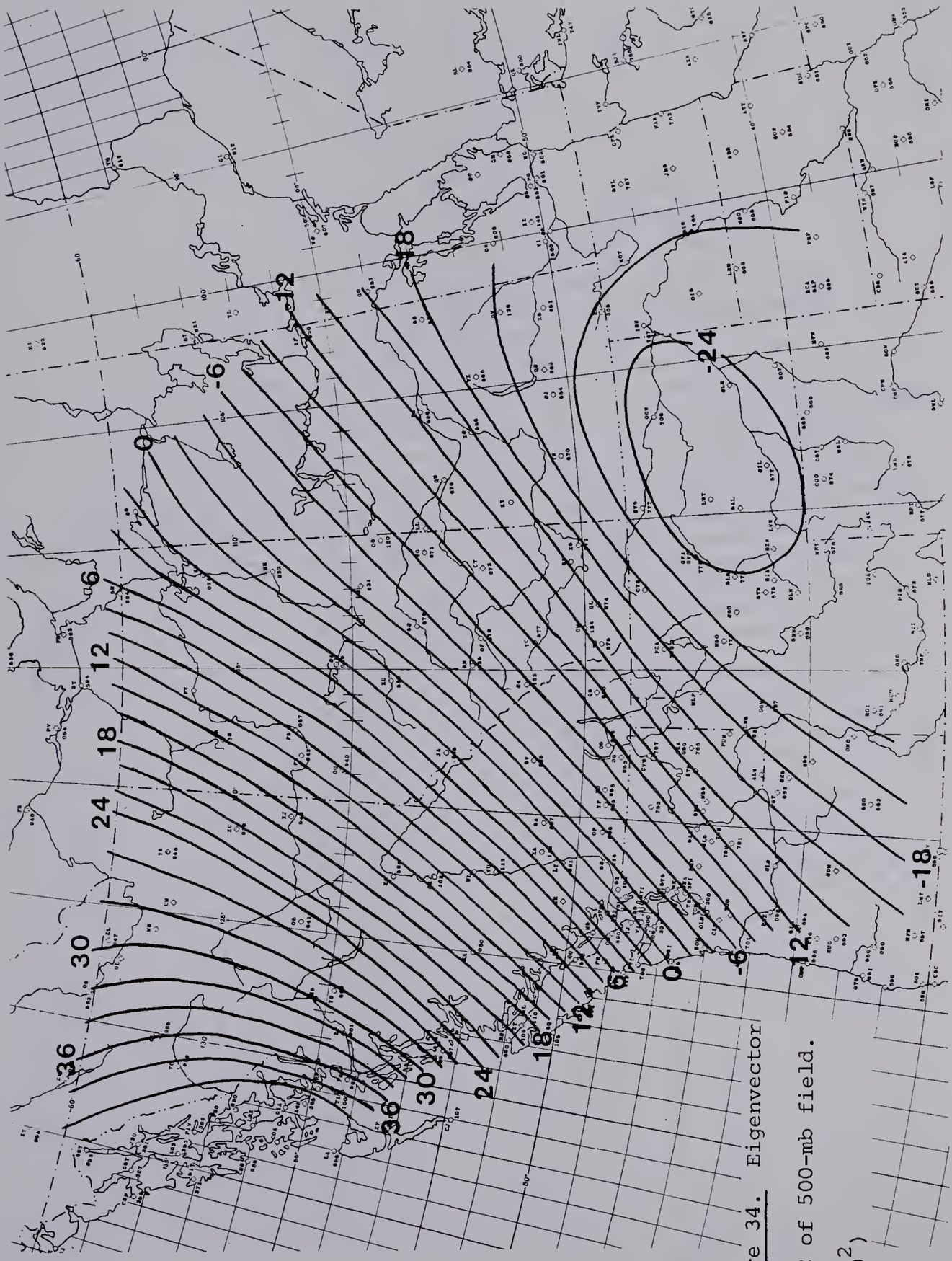

Figure 34. Eigenvector No. 2 of 500-mb field. ($\times 10^2$)

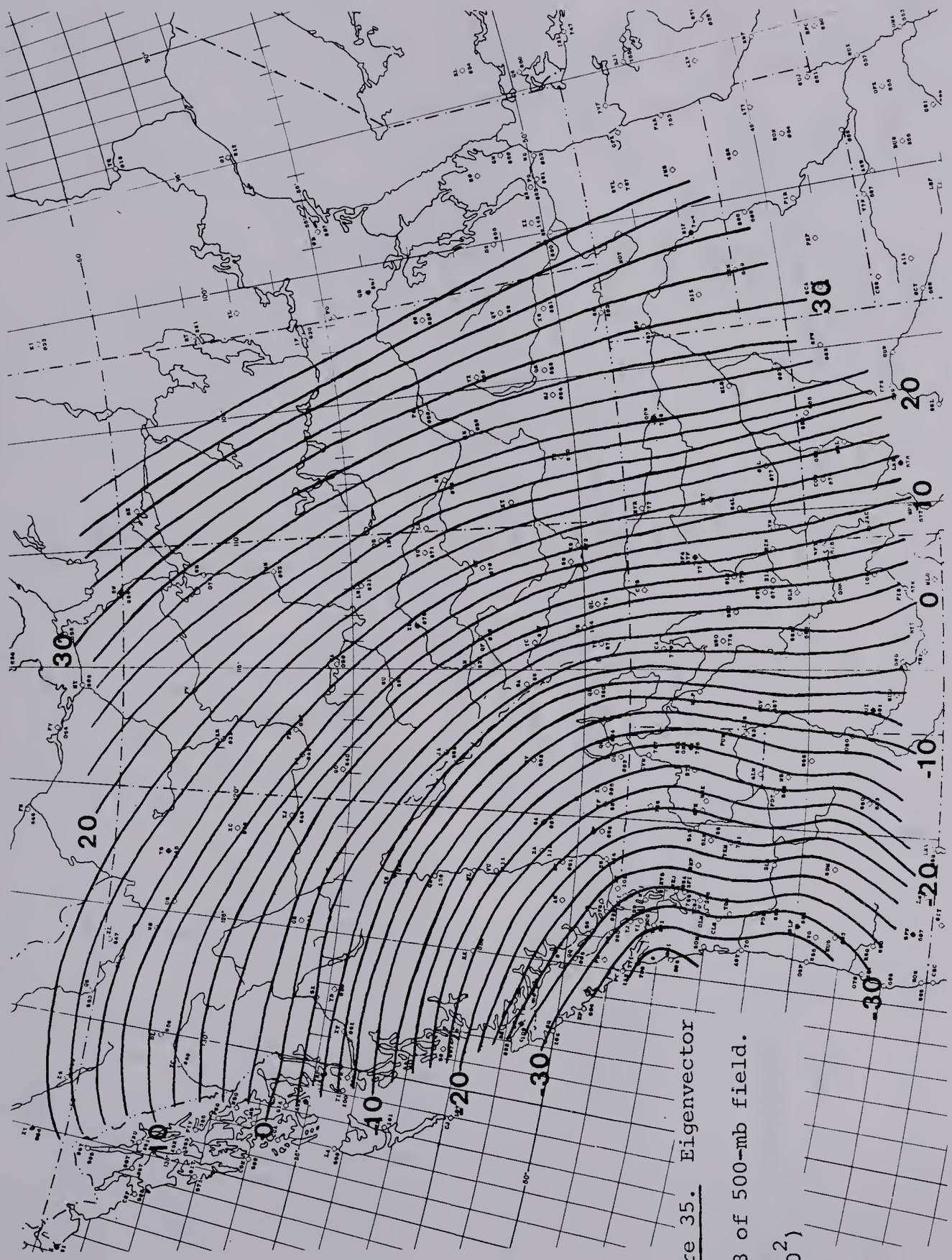

Figure 35. Eigenvector No. 3 of 500-mb field. ($\times 10^2$)

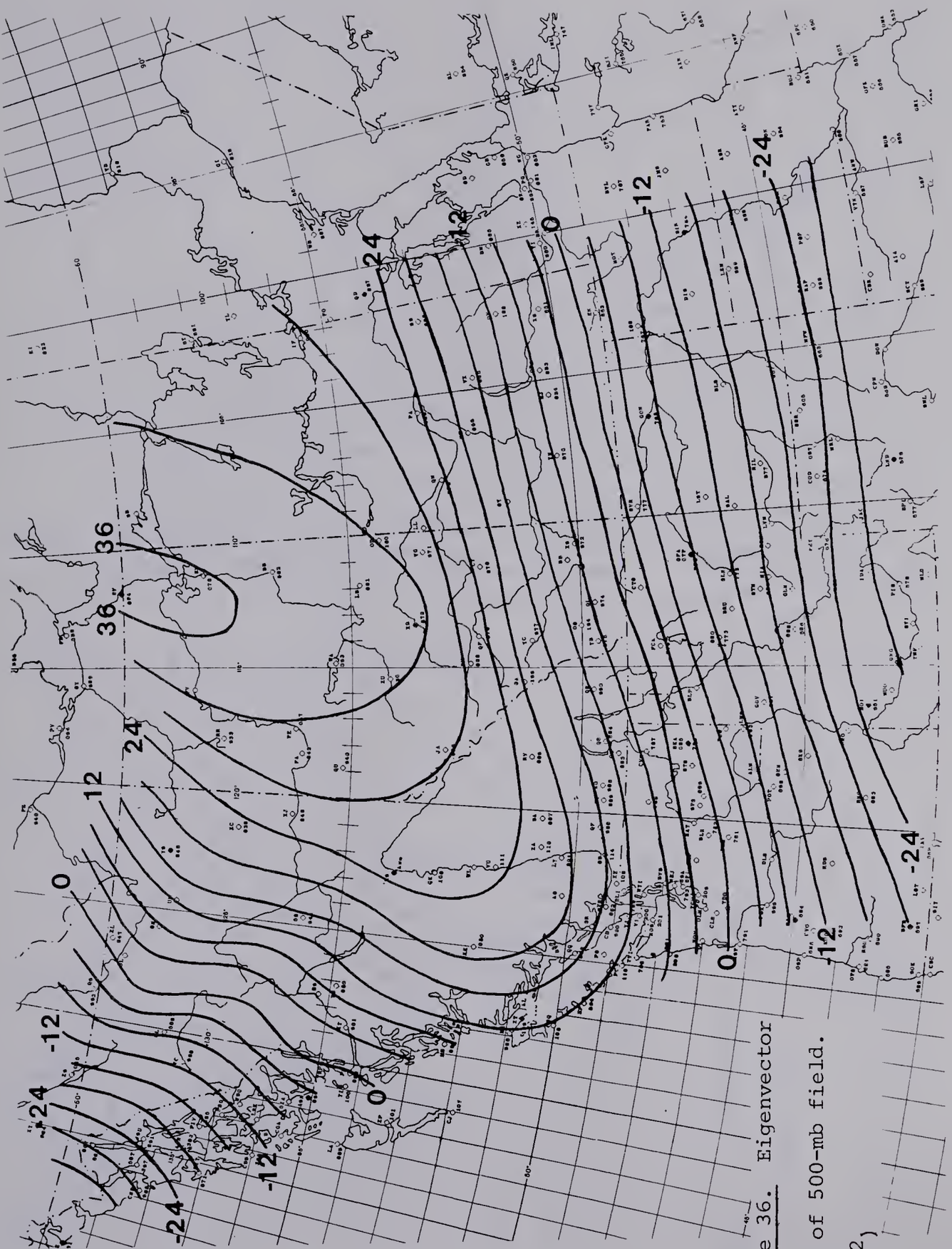

Figure 36. Eigenvector No. 4 of 500-mb field. ($\times 10^2$)

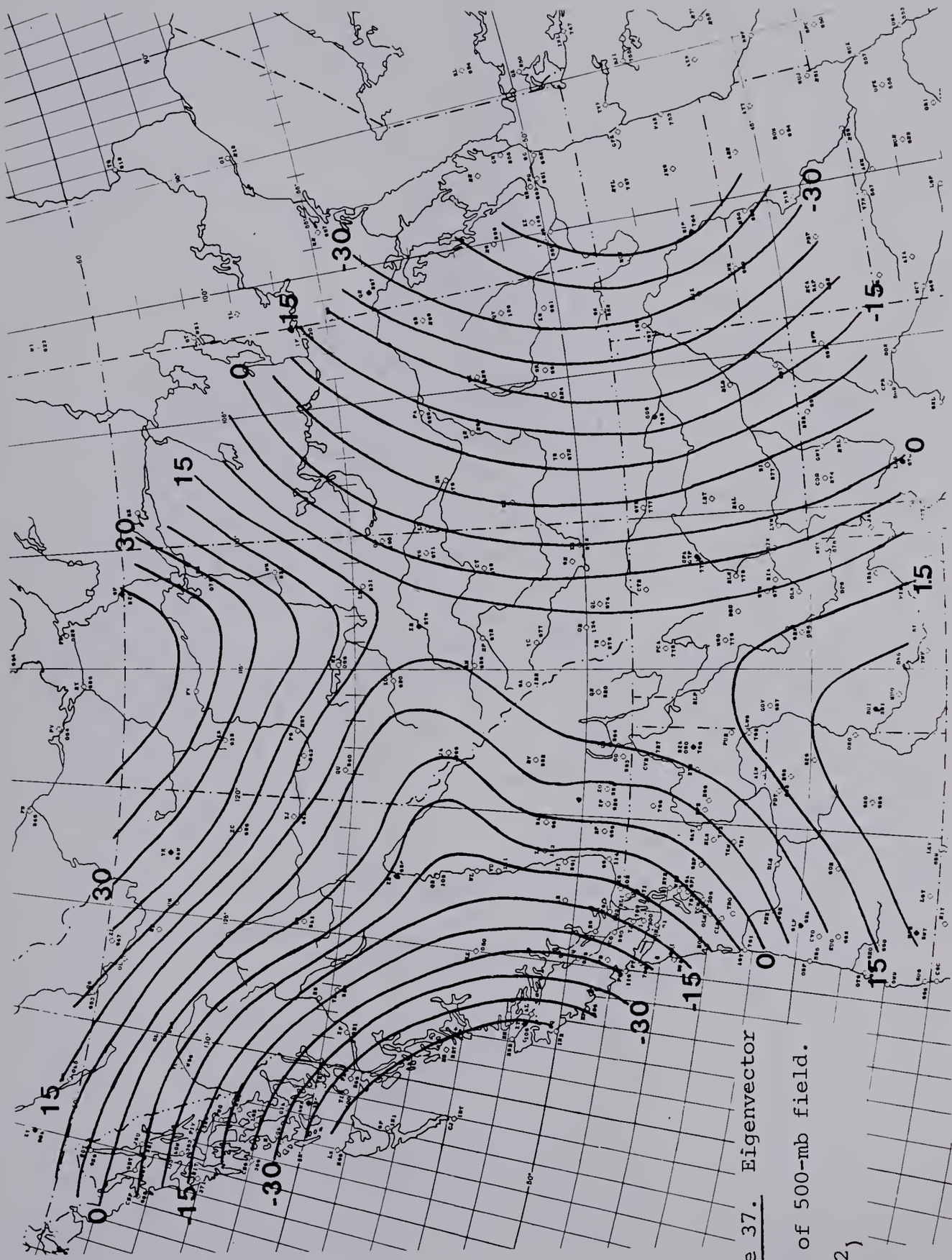

Figure 37. Eigenvector No. 5 of 500-mb field. ($\times 10^2$)

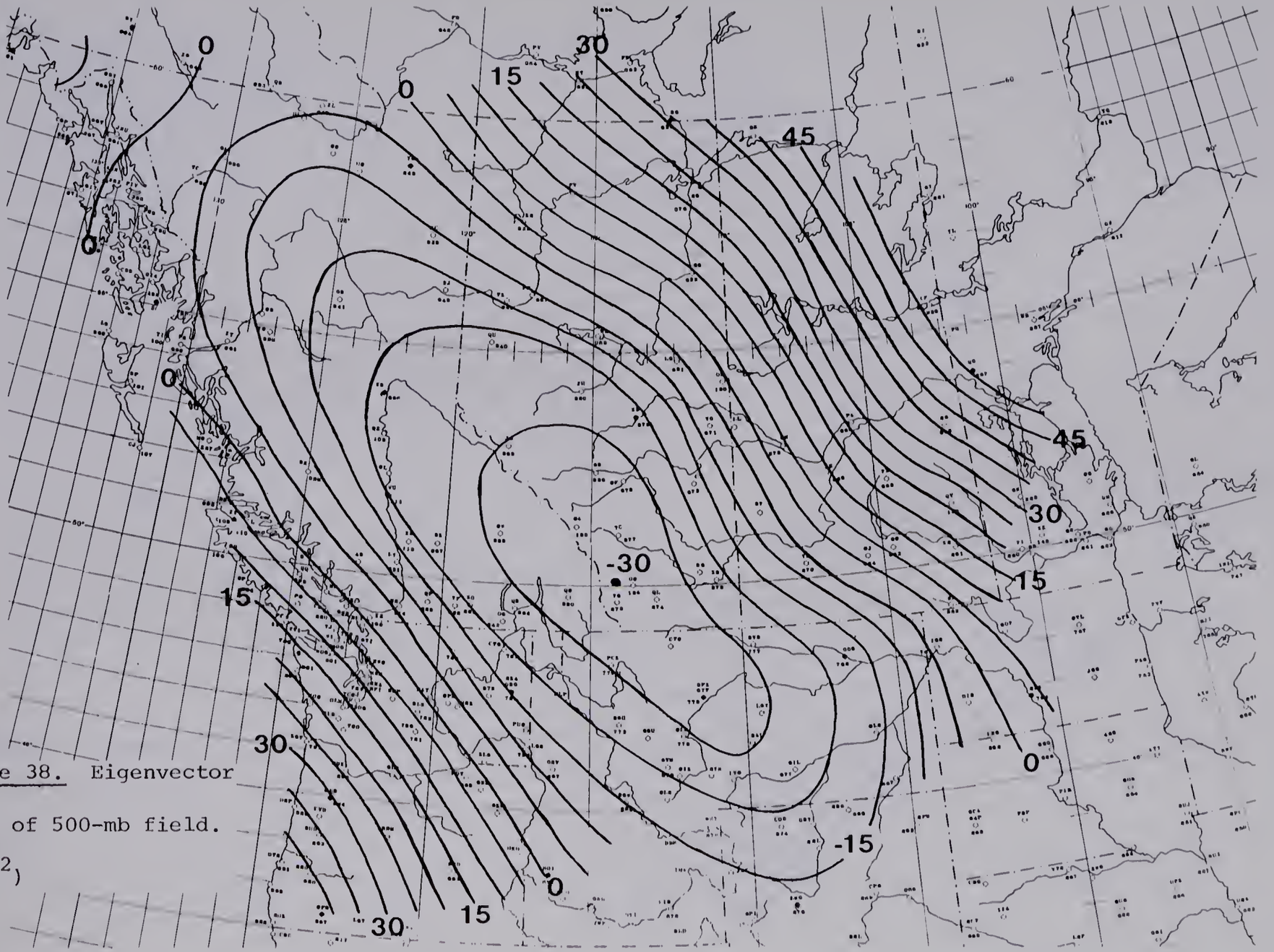

Figure 38. Eigenvector No. 6 of 500-mb field. (x 10^2)

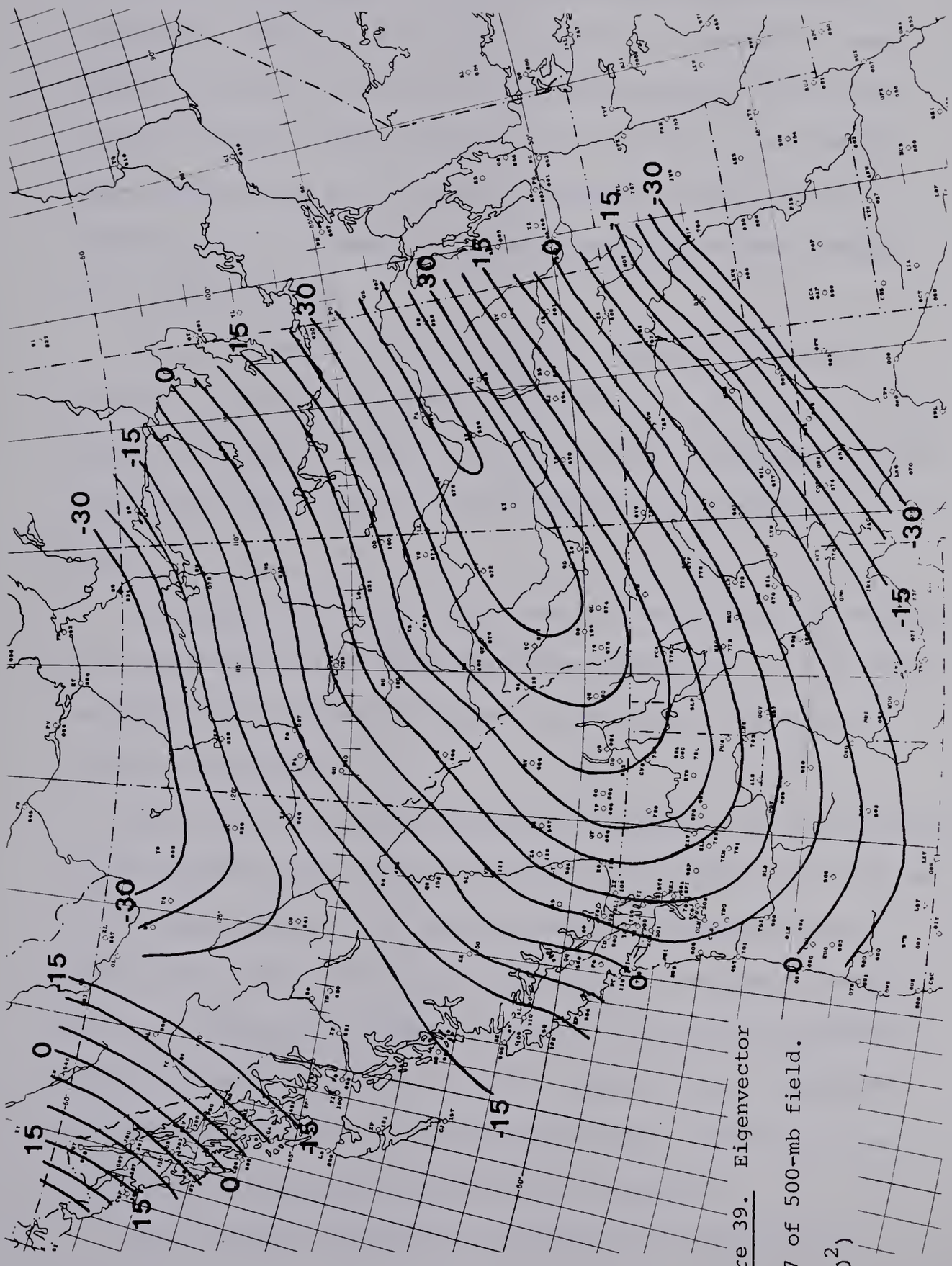

Figure 39. Eigenvector No. 7 of 500-mb field. (x 10^2)

The time coefficients in themselves were not particularly interesting. However, an attempt was made to discover any wave-type phenomena as in Section 3.2 by cross-correlating the first ten time functions using lags from -30 to +30 days. The results were disappointing, with a maximum correlation coefficient of approximately 0.3. Hence no further investigations were carried out in this direction.

The only practical use of empirical orthogonal functions mentioned so far was the possibility of economically storing climatological variables. Another use of these functions, pointed out by Craddock and Flood (1969), concerns the detection of errors in the data.

As suggested by Craddock and Flood the coefficients of kurtosis were calculated for the first eight time coefficients. Even with a few bogus values placed in the data no suspicious values for the coefficient resulted.

Next, the first eight eigenvectors, accounting for 97.8 per cent of the variance, were combined with their time coefficients and the results were compared with the input data. Differences greater than 50 gpm. were checked against the original radiosonde reports and of the 390 values checked, 21 errors were found and corrected. Combining nine eigenvectors and their time functions, as derived from the corrected data, 290 differences greater than 50 gpm. were found with 13 errors being detected and corrected.

Using the same eigenvectors as calculated from the years

Eigenvector Number	Var. Explained (with Errors)	Change After Errors Removed
1	57.398	+0.043
2	16.267	+0.016
3	12.205	+0.003
4	5.460	+0.005
5	2.671	+0.007
6	2.068	-0.003
7	1.142	-0.003
8	0.653	-0.003
9	0.476	-0.009
10	0.358	-0.003
11	0.276	-0.007
12	0.266	-0.003
13	0.154	-0.010
14	0.150	-0.006
15	0.129	-0.005
16	0.105	-0.007
17	0.086	-0.005
18	0.081	-0.002
19	0.057	-0.005

Table 8. Change in the variance explained by successive eigenvectors after removal of errors from the 500-mb field.

1963 to 1972, the time coefficients for an independent sample from 1973 were computed. Utilizing the first eight eigenvectors to check for errors, nine were found out of a total of 57 differences greater than 50 gpm.

An interesting result of the removal of the errors was that the amount of variance explained by the lower-numbered eigenvectors increased while that explained by the higher-numbered eigenvectors decreased as shown in Table 8. The differences, though small, support the contention that the higher-numbered eigenvectors are most likely nothing but noise. There were also slight changes in the eigenvectors themselves, ranging from 0.005 per cent for

eigenvector No. 1 to over 6 per cent for number 8.

Empirical orthogonal functions thus appear a good tool for error detection in some of the data but, unfortunately, they have their limitations. Even with over 97.8 per cent of the variance explained, over 10 per cent of the recombined values had differences of about 10 per cent from the sample values. Certainly some of these differences were introduced by instrument malfunction when recording the raw data, and in the subsequent handling of the data, but some of these differences must have been caused by real processes.

This was evidenced by studying the difference fields and noting that, in some places, the errors tended to clump in groups. By checking 500-mb maps on the corresponding days it was found that a sudden low or small trough moved through the northern United States. This further points out the problem of where to draw the line in an eigenvector representation, since to discard eigenvectors higher than number nine would have meant filtering out a real phenomenon. Perhaps one of the higher-numbered eigenvectors could have helped to explain most of the differences caused by these sudden lows.

Two other investigations were carried out, using only one value per day, and using only two years of data. Utilizing one value per day, the first eight eigenvectors stayed the same with approximately the same amount of variance explained by each, indicating perhaps that the first eight represent a time

Eigenvector Number	Per Cent Variance	Cumulative Variance
1	26.0	26.0
2	15.7	41.7
3	7.9	49.6
4	5.8	55.3
5	5.7	61.1
6	5.1	66.1
7	4.5	70.6
8	3.4	74.0
9	2.9	77.0
10	2.7	79.7
11	2.5	82.2
12	2.4	84.6
13	2.1	86.6
14	2.0	88.6
15	1.8	90.5
16	1.6	92.1
17	1.6	93.6
18	1.5	95.1
19	1.2	96.3
20	1.1	97.4
21	1.1	98.5
22	0.9	99.4
23	0.6	100.0

Table 9. Variances and cumulative variances explained by successive eigenvectors of the precipitation data.

scale greater than 12 hours. Using two years of data resulted in the first four eigenvector patterns being almost identical to the first four derived from ten years of data, indicating that the first four represent some basic deviations from the mean inherent in the 500-mb field during the summer.

3.5 Analysis of Precipitation

Using 23 stations almost entirely in Alberta, eigenanalysis was performed on ten years of 24-hourly precipitation data cor-

responding to the dates of the analysis of the 500-mb heights. Table 9 shows the variances explained by the various eigenvectors.

Day-to-day variations in precipitation tend to be very erratic. This was evident in the low amount of variance explained by the lower-numbered eigenvectors as well as by the irregular time coefficients. Detailed analysis might have produced some useful results, but such further work was not carried out.

CHAPTER 4

CONCLUSION

4.1 Comments and Considerations for Future Work

Eigenanalysis is a means of recording data with relatively large savings in the space required for storage. Problems arise because of the possibility of filtering out not only noise (which is desirable) but also real physical occurrences. Before any use is made of eigenvectors for this purpose, more work must be carried out on where the line should be drawn between noise and real data.

One corollary of the above problem is that if the higher numbered eigenvectors do not represent random variations or errors, then by studying these patterns an insight may be gained into unusual or rare phenomena.

A problem arises also with the first eigenvector pattern and its similarity to the variance map, except in the one case of the 500-mb field. If the data are analyzed without removing the means, the first eigenvector will represent the mean field. If the means are removed and the variances are not removed through normalization, then the first eigenvector will resemble the variance field. If two or more different variables are to be analyzed then normalization is a must. It is the opinion of this author that normalization should always be carried out when eigenanalysis is performed no

matter how many variables are analyzed.

Error detection played an important role in this paper and empirical orthogonal functions showed a real application in the detection of false values in the 500-mb heights. Out of a total of about 600 possible errors checked, 34 faulty values were discovered. Whether or not these were all the errors in the data is difficult to determine since there may have been erroneous values in the original data and also the 3060 values were not verified; but if there were only 34 errors, then verifying 600 values seems much easier than verifying over 3000.

An understanding of the eigenvector patterns is necessary in order to lend some "reality" when working with them. Interpretation has always been a problem and one possible investigation which may elucidate their meaning would be to combine each eigenvector separately with the mean field and display the results on a screen as a "movie". The cumulative patterns could also be displayed; the mean field plus eigenvector No. 1, then the mean field plus eigenvector No. 1 plus eigenvector No. 2, etc.

Because data collecting points are seldom on an evenly spaced grid, empirical orthogonal functions ease analysis since the spatial distribution of stations need not be regular. This could also be important where a large-scale flow is required, say over mountains, with studies being conducted in a small area on the lee side.

If many values in a long series of observations of a fairly smooth variable (not as discontinuous as precipitation) are missing, eigenanalysis should prove one of the best methods of filling in

these spaces, since the values computed will be derived from the past and future history of the variable, not climatology, persistence, etc. Empirical orthogonal functions may also prove to be efficient in interpolating and extrapolating variables both in time and space.

In the solution of differential equations where the grid does not encompass the earth but rather where there are boundary conditions to contend with, eigenanalysis may offer a better scheme for filling in these values than, say, assuming steady state conditions. For example, if the equations of a weather prediction model were being solved numerically on a grid encompassing Alberta and one of the parameters was the 500-mb field, then using the mean field plus eigenvector No. 1 would give a more realistic picture of the pattern at the boundaries than, say, the mean field alone.

The brightness of clouds as obtained from satellite images is affected across each scan by both the angle of the sun from the zenith and the position of the satellite, as well as whether the satellite is scanning towards the sun or away from it. The problem of correcting for this has so far not been solved satisfactorily, but if many satellite images were to undergo eigenanalysis, then possibly a matrix of correction values could be obtained in order to compensate for these differences in brightness.

One of the most promising developments in the use of empirical orthogonal functions appears to be in complex eigenvector analysis as shown by Wallace and Dickinson (1972). By utilizing the cross-spectrum matrix instead of the covariance or cross-correlation matrix information is obtained about both the phase angle and the

amplitude of the correlation between different variables. One use of complex eigenvector analysis was shown by Wallace (1972) and many other applications are feasible such as the analysis of micro-meteorological variables. Since cross-spectrum analysis is often performed on data observed close to the ground (and these measurements may involve variables such as amount of pollutants), complex eigenvector analysis seems ideal in aiding in the study of the correlations, both in amplitude and phase angle, of these variables.

4.2 Concluding Remarks

Empirical orthogonal functions have seen little use in meteorology primarily because eigenanalysis is relatively unknown. Also, in the study of different variables and how they interact, dynamic procedures are favored over statistical ones. The equations of meteorology are complex and at present unsolvable analytically; not even all the physics of the atmosphere is clearly understood. Statistical methods have the advantage of taking a set of numbers representing variables and analyzing them in a way that perhaps sheds some light on the actual dynamic processes, thus paving the way for comprehending some of the phenomena which are not now apparent.

The use of eigenvectors and their associated time coefficients may be just such a statistical method for unravelling some of the interrelationships of different meteorological variables though much more basic work must be done in the understanding of these creatures.

REFERENCES

Lorenz, E.N., 1956: Empirical orthogonal functions and statistical weather prediction. Sci. Rept. No. 1, Contract AF19(604)-1566, Dept. of Meteorology, M.I.T., 49pp.

Gilman, D.L., 1957: Empirical orthogonal functions applied to thirty-day forecasting. Sci. Rept., No. 1, Contract AF19(604)-1283, Dept. of Meteorology, M.I.T., 129 pp.

Craddock, J.M. and Flood, C.R., 1969: Eigenvectors for representing the 500-mb geopotential surface over the Northern Hemisphere, Quart., J.R. Met. Soc., 95, pp. 576-593

Craddock, J.M. and Flintoff, S., 1970: Eigenvector representations of Northern Hemisphere fields, Quart., J.R. Met. Soc., 96, pp. 124-129

Kutzbach, J.E., 1967: Empirical Eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America, J. Appl. Meteor., 5, pp. 791-802

Wallace, J.M. and Dickinson, R.E., 1972: Empirical orthogonal representation of time series in the frequency domain. Part I: Theoretical considerations, J. Appl. Meteor., 6, pp. 887-892

Wallace, J.M., 1972: Empirical orthogonal representation of time series in the frequency domain. Part II: Application to the study of tropical wave disturbances, J. Appl. Meteor., 6, pp. 893-900

APPENDIX I

A BRIEF SUMMARY OF

LORENZ'S DERIVATION OF EMPIRICAL ORTHOGONAL FUNCTIONS

If the set of M variables to be analyzed is represented by $p_1(t),\ldots p_M(t)$, observed at the N times $t_1,\ldots t_N$ then the total variance of the variables can be written as:

$$V=\sum_{m=1}^{M} \overline{p_m^{*2}} \qquad \ldots 1$$

where the star indicates departure from the mean.

Let $q_1(t),\ldots q_K(t)$ be any K quantities such that K is less than M and let

$$p_m^*(t_i)=\sum_{k=1}^{K} y_{km}q_k(t_i)+r_m(t_i) \qquad \ldots 2$$

where the y_{km} are chosen to minimize the error $r_m(t_i)$ for each m and hence to minimize the total "unexplained" variance R,

$$R=\sum_{m=1}^{M} \overline{r_m^2} \qquad \ldots 3$$

Once the y_{km} are chosen, the q_k have to be picked so that the minimum value of R is minimized and the quantity (V-R)/V becomes the fraction of the total variance which can be represented by the K quantities.

APPENDIX II

AN INTRODUCTION TO PERIODOGRAMS

Any variable p(t) can be expressed as a Fourier series as follows:

$$p(t)=a_0/2 + \sum_{n=1}^{\infty} (a_n\cos(2n\pi t/T) + b_n\sin(2n\pi t/T)) \quad ...1$$

with $0 < t < T$

where T is the length of sampling.

If p(t) is sampled at uniform intervals t_i then equation 1 becomes:

$$p(t)=a_0/2 + \sum_{n=1}^{N} (a_n\cos(2\pi nt/T) + b_n\sin(2\pi nt/T)) \quad ...2$$

with $N=T/(t_{i+1}-t_i)$,

and the coefficients a_n and b_n are determined by,

$$a_n=(2/T)\sum_{i=1}^{N} p(t_i)\cos(2\pi nt_i/T) \quad ...3$$

$$b_n=(2/T)\sum_{i=1}^{N} p(t_i)\sin(2\pi nt_i/T) \quad ...4$$

The plot of $R_n^2 = a_n^2 + b_n^2$ versus the period T/n is called a periodogram and gives an indication of the contribution of each period T/n to the total variance of p(t).

Now let

$$p_m^*(t)=\sum_{k=1}^{K} Y_{km}Q_k^*(t) \qquad \ldots 4$$

where Y_{km} and $Q_k^*(t)$ are chosen such that

$$\sum_{m=1}^{M} Y_{km}Y_{jm} = \delta_{kj} \qquad \ldots 5$$

and

$$N\overline{Q_k^*Q_j^*} = a_k\delta_{kj} \qquad \ldots 6$$

where $a_k \geq a_{k+1} \geq 0$.

These Q*'s satisfy the requirements for the q's and it can also be shown that

$$V=(1/N)\sum_{k=1}^{M} a_k \qquad \ldots 7$$

and

$$V-R=(1/N)\sum_{k=1}^{K} a_k.$$

The Q*'s are the same as those in Chapter 1 while the Y's are the same as the X's. Proofs of these results can be found in Lorenz's paper. (Lorenz, 1956)

APPENDIX III

A LISTING OF THE SUBROUTINES USED TO FIND

THE EIGENVECTORS OF A REAL, SYMMETRIC MATRIX

```
C
C               SUBROUTINE TRED2(NM,N,A,D,E,Z)
C     ------------------------------------------------------------
C
C
C     THIS SUBROUTINE IS A TRANSLATION OF THE ALGOL PROCEDURE TRED2,
C     NUM. MATH. 11, 181-195(1968) BY MARTIN, REINSCH, AND WILKINSON.
C     HANDBOOK FOR AUTO. COMP., VOL.II-LINEAR ALGEBRA, 212-226(1971).
C
C     THIS SUBROUTINE REDUCES A REAL SYMMETRIC MATRIX TO A
C     SYMMETRIC TRIDIAGONAL MATRIX USING AND ACCUMULATING
C     ORTHOGONAL SIMILARITY TRANSFORMATIONS.
C
C     ON INPUT-
C
C        NM MUST BE SET TO THE ROW DIMENSION OF TWO-DIMENSIONAL
C          ARRAY PARAMETERS AS DECLARED IN THE CALLING PROGRAM
C          DIMENSION STATEMENT,
C
C        N IS THE ORDER OF THE MATRIX,
C
C        A CONTAINS THE REAL SYMMETRIC INPUT MATRIX.  ONLY THE
C          LOWER TRIANGLE OF THE MATRIX NEED BE SUPPLIED.
C
C     ON OUTPUT-
C
C        D CONTAINS THE DIAGONAL ELEMENTS OF THE TRIDIAGONAL MATRIX,
C
C        E CONTAINS THE SUBDIAGONAL ELEMENTS OF THE TRIDIAGONAL
C          MATRIX IN ITS LAST N-1 POSITIONS.  E(1) IS SET TO ZERO,
C
C        Z CONTAINS THE ORTHOGONAL TRANSFORMATION MATRIX
```

```
C           PRODUCED IN THE REDUCTION,
C
C        A AND Z MAY COINCIDE.  IF DISTINCT, A IS UNALTERED.
C
C     QUESTIONS AND COMMENTS SHOULD BE DIRECTED TO B. S. GARBOW,
C     APPLIED MATHEMATICS DIVISION, ARGONNE NATIONAL LABORATORY
C
C     ------------------------------------------------------------
C
      SUBROUTINE TRED2(NM,N,A,D,E,Z)
C
      INTEGER I,J,K,L,N,II,NM,JP1
      REAL A(NM,N),D(N),E(N),Z(NM,N)
      REAL F,G,H,HH,SCALE
      REAL SQRT,AES,SIGN
C
C
      DO 100 I = 1, N
C
         DO 100 J = 1, I
            Z(I,J) = A(I,J)
  100 CONTINUE
C
      IF (N .EQ. 1) GO TO 320
C     ********** FOR I=N STEP -1 UNTIL 2 DO -- **********
      DO 300 II = 2, N
         I = N + 2 - II
         L = I - 1
         H = 0.0
         SCALE = 0.0
         IF (L .LT. 2) GO TO 130
C     ********** SCALE ROW (ALGOL TOL THEN NOT NEEDED) **********
```

```
            DO 120 K = 1, L
  120       SCALE = SCALE + ABS(Z(I,K))
C
            IF (SCALE .NE. 0.0) GO TO 140
  130       E(I) = Z(I,L)
            GO TO 290
C
  140       DO 150 K = 1, L
               Z(I,K) = Z(I,K) / SCALE
               H = H + Z(I,K) * Z(I,K)
  150       CONTINUE
C
            F = Z(I,L)
            G = -SIGN(SQRT(H),F)
            E(I) = SCALE * G
            H = H - F * G
            Z(I,L) = F - G
            F = 0.0
C
            DO 240 J = 1, L
               Z(J,I) = Z(I,J) / (SCALE * H)
               G = 0.0
C        ********** FORM ELEMENT OF A*U **********
               DO 180 K = 1, J
  180          G = G + Z(J,K) * Z(I,K)
C
               JP1 = J + 1
               IF (L .LT. JP1) GO TO 220
C
               DO 200 K = JP1, L
  200          G = G + Z(K,J) * Z(I,K)
C        ********** FORM ELEMENT OF P **********
```

```
  220        E(J) = G / H
             F = F + E(J) * Z(I,J)
  240     CONTINUE
C
          HH = F / (H + H)
C      ********** FORM REDUCED A **********
          DO 260 J = 1, L
             F = Z(I,J)
             G = E(J) - HH * F
             E(J) = G
C
             DO 260 K = 1, J
                Z(J,K) = Z(J,K) - F * E(K) - G * Z(I,K)
  260     CONTINUE
C
          DO 280 K = 1, L
  280     Z(I,K) = SCALE * Z(I,K)
C
  290     D(I) = H
  300  CONTINUE
C
  320  D(1) = 0.0
       E(1) = 0.0
C      ********** ACCUMULATION OF TRANSFORMATION MATRICES **********
       DO 500 I = 1, N
          L = I - 1
          IF (D(I) .EQ. 0.0) GO TO 380
C
          DO 360 J = 1, L
             G = 0.0
C
             DO 340 K = 1, L
```

```
  340         G = G + Z(I,K) * Z(K,J)
C
              DO 360 K = 1, L
                 Z(K,J) = Z(K,J) - G * Z(K,I)
  360       CONTINUE
C
  380       D(I) = Z(I,I)
            Z(I,I) = 1.0
            IF (L .LT. 1) GO TO 500
C
            DO 400 J = 1, L
               Z(I,J) = 0.0
               Z(J,I) = 0.0
  400       CONTINUE
C
  500 CONTINUE
C
      RETURN
C     ********** LAST CARD OF TRED2 **********
      END
```

```
C
C                SUBROUTINE TQL2(NM,N,D,E,Z,IERR)
C     ------------------------------------------------------------
C
C
C     THIS SUBROUTINE IS A TRANSLATION OF THE ALGOL PROCEDURE TQL2,
C     NUM. MATH. 11, 293-306(1968) BY BOWDLER, MARTIN, REINSCH, AND
C     WILKINSON.
C     HANDBOOK FOR AUTO. COMP., VOL.II-LINEAR ALGEBRA, 227-240(1971).
C
C     THIS SUBROUTINE FINDS THE EIGENVALUES AND EIGENVECTORS
C     OF A SYMMETRIC TRIDIAGONAL MATRIX BY THE QL METHOD.
C     THE EIGENVECTORS OF A FULL SYMMETRIC MATRIX CAN ALSO
C     BE FOUND IF  TRED2  HAS BEEN USED TO REDUCE THIS
C     FULL MATRIX TO TRIDIAGONAL FORM.
C
C     ON INPUT-
C
C        NM MUST BE SET TO THE ROW DIMENSION OF TWO-DIMENSIONAL
C          ARRAY PARAMETERS AS DECLARED IN THE CALLING PROGRAM
C          DIMENSION STATEMENT,
C
C        N IS THE ORDER OF THE MATRIX,
C
C        D CONTAINS THE DIAGONAL ELEMENTS OF THE INPUT MATRIX,
C
C        E CONTAINS THE SUBDIAGONAL ELEMENTS OF THE INPUT MATRIX
C          IN ITS LAST N-1 POSITIONS.  E(1) IS ARBITRARY,
C
C        Z CONTAINS THE TRANSFORMATION MATRIX PRODUCED IN THE
C          REDUCTION BY  TRED2, IF PERFORMED.  IF THE EIGENVECTORS
C          OF THE TRIDIAGONAL MATRIX ARE DESIRED, Z MUST CONTAIN
```

```
C           THE IDENTITY MATRIX.
C
C        ON OUTPUT-
C
C           D CONTAINS THE EIGENVALUES IN ASCENDING ORDER.  IF AN
C             ERROR EXIT IS MADE, THE EIGENVALUES ARE CORRECT BUT
C             UNORDERED FOR INDICES 1,2,...,IERR-1,
C
C           E HAS BEEN DESTROYED,
C
C           Z CONTAINS ORTHONORMAL EIGENVECTORS OF THE SYMMETRIC
C             TRIDIAGONAL (OR FULL) MATRIX.  IF AN ERROR EXIT IS MADE,
C             Z CONTAINS THE EIGENVECTORS ASSOCIATED WITH THE STORED
C             EIGENVALUES,
C
C           IERR IS SET TO
C             ZERO       FOR NORMAL RETURN,
C             J          IF THE J-TH EIGENVALUE HAS NOT BEEN
C                        DETERMINED AFTER 30 ITERATIONS.
C
C     QUESTIONS AND COMMENTS SHOULD BE DIRECTED TO B. S. GARBOW,
C     APPLIED MATHEMATICS DIVISION, ARGONNE NATIONAL LABORATORY
C
C     ------------------------------------------------------------------
C
      SUBROUTINE TQL2(NM,N,D,E,Z,IERR)
C
      INTEGER I,J,K,L,M,N,II,L1,NM,MML,IERR
      REAL D(N),E(N),Z(NM,N)
      REAL B,C,F,G,H,P,R,S,MACHEP
C     REAL SQRT,ABS,SIGN
C
```

```
C
C      ********** MACHEP IS A MACHINE DEPENDENT PARAMETER SPECIFYING
C                 THE RELATIVE PRECISION OF FLOATING POINT ARITHMETIC.
C
C                 **********
C
       MACHEP=.9536743E-06
C
       IERR = 0
       IF (N .EQ. 1) GO TO 1001
C
       DO 100 I = 2, N
  100  E(I-1) = E(I)
C
       F = 0.0
       B = 0.0
       E(N) = 0.0
C
       DO 240 L = 1, N
          J = 0
          H = MACHEP * (ABS(D(L)) + ABS(E(L)))
          IF (B .LT. H) B = H
C      ********** LOOK FOR SMALL SUB-DIAGONAL ELEMENT **********
          DO 110 M = L, N
             IF (ABS(E(M)) .LE. B) GO TO 120
C      ********** E(N) IS ALWAYS ZERO, SO THERE IS NO EXIT
C                 THROUGH THE BOTTOM OF THE LOOP **********
  110     CONTINUE
C
  120     IF (M .EQ. L) GO TO 220
  130     IF (J .EQ. 30) GO TO 1000
          J = J + 1
```

```
C      ********** FORM SHIFT **********
          L1 = L + 1
          G = D(L)
          P = (D(L1) - G) / (2.0 * E(L))
          R = SQRT(P*P+1.0)
          D(L) = E(L) / (P + SIGN(R,P))
          H = G - D(L)
C
          DO 140 I = L1, N
  140     D(I) = D(I) - H
C
          F = F + H
C      ********** QL TRANSFORMATION **********
          P = D(M)
          C = 1.0
          S = 0.0
          MML = M - L
C      ********** FOR I=M-1 STEP -1 UNTIL L DO -- **********
          DO 200 II = 1, MML
             I = M - II
             G = C * E(I)
             H = C * P
             IF (ABS(P) .LT. ABS(E(I))) GO TO 150
             C = E(I) / P
             R = SQRT(C*C+1.0)
             E(I+1) = S * P * R
             S = C / R
             C = 1.0 / R
             GO TO 160
  150        C = P / E(I)
             R = SQRT(C*C+1.0)
             E(I+1) = S * E(I) * R
```

```
            S = 1.0 / H
            C = C * S
  160       P = C * D(I) - S * G
            D(I+1) = H + S * (C * G + S * D(I))
C     ********** FORM VECTOR **********
            DO 180 K = 1, N
               H = Z(K,I+1)
               Z(K,I+1) = S * Z(K,I) + C * H
               Z(K,I) = C * Z(K,I) - S * H
  180       CONTINUE
C
  200    CONTINUE
C
         E(L) = S * P
         D(L) = C * P
         IF (ABS(E(L)) .GT. B) GO TO 130
  220    D(L) = D(L) + F
  240 CONTINUE
C     ********** ORDER EIGENVALUES AND EIGENVECTORS **********
      DO 300 II = 2, N
         I = II - 1
         K = I
         P = D(I)
C
         DO 260 J = II, N
            IF (D(J) .GE. P) GO TO 260
            K = J
            P = D(J)
  260    CONTINUE
C
         IF (K .EQ. I) GO TO 300
         D(K) = D(I)
```

```
            D(I) = P
C
            DO 280 J = 1, N
               P = Z(J,I)
               Z(J,I) = Z(J,K)
               Z(J,K) = P
  280       CONTINUE
C
  300 CONTINUE
C
      GO TO 1001
C     ********** SET ERROR -- NO CONVERGENCE TO AN
C                EIGENVALUE AFTER 30 ITERATIONS **********
 1000 IERR = L
 1001 RETURN
C     ********** LAST CARD OF TQL2 **********
      END
```